



City Research Online

City, University of London Institutional Repository

Citation: Hagen, S., Bugge, C., Dean, S., Elders, A., Hay-Smith, J., Kilonzo, M., McClurg, D., Abdel-Fattah, M., Agur, W., Andreis, F., et al (2020). Basic versus biofeedback-mediated intensive pelvic floor muscle training for women with urinary incontinence: the OPAL RCT. Health Technology Assessment, 24(70), doi: 10.3310/hta24700

This is the published version of the paper.

This version of the publication may differ from the final published version.

Permanent repository link: <https://openaccess.city.ac.uk/id/eprint/25985/>

Link to published version: <https://doi.org/10.3310/hta24700>

Copyright: City Research Online aims to make research outputs of City, University of London available to a wider audience. Copyright and Moral Rights remain with the author(s) and/or copyright holders. URLs from City Research Online may be freely distributed and linked to.

Reuse: Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge. Provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

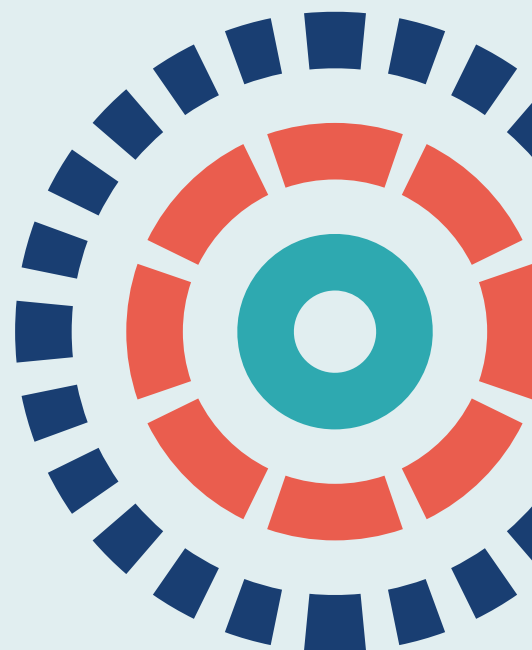
Health Technology Assessment

Volume 24 • Issue 70 • December 2020

ISSN 1366-5278

Basic versus biofeedback-mediated intensive pelvic floor muscle training for women with urinary incontinence: the OPAL RCT

*Suzanne Hagen, Carol Bugge, Sarah G Dean, Andrew Elders, Jean Hay-Smith,
Mary Kilonzo, Doreen McClurg, Mohamed Abdel-Fattah, Wael Agur, Federico Andreis,
Joanne Booth, Maria Dimitrova, Nicola Gillespie, Cathryn Glazener, Aileen Grant,
Karen L Guerrero, Lorna Henderson, Marija Kovandzic, Alison McDonald, John Norrie,
Nicole Sergenson, Susan Stratton, Anne Taylor and Louise R Williams*



Basic versus biofeedback-mediated intensive pelvic floor muscle training for women with urinary incontinence: the OPAL RCT

Suzanne Hagen,^{1*} Carol Bugge,² Sarah G Dean,³ Andrew Elders,¹ Jean Hay-Smith,⁴ Mary Kilonzo,⁵ Doreen McClurg,¹ Mohamed Abdel-Fattah,⁵ Wael Agur,⁶ Federico Andreis,² Joanne Booth,⁷ Maria Dimitrova,⁵ Nicola Gillespie,¹ Cathryn Glazener,⁸ Aileen Grant,⁹ Karen L Guerrero,¹⁰ Lorna Henderson,¹¹ Marija Kovandzic,² Alison McDonald,¹¹ John Norrie,¹² Nicole Sergenson,¹ Susan Stratton,¹ Anne Taylor² and Louise R Williams¹

¹Nursing, Midwifery and Allied Health Professions Research Unit, Glasgow Caledonian University, Glasgow, UK

²Faculty of Health Sciences and Sport, University of Stirling, Stirling, UK

³University of Exeter Medical School, Exeter, UK

⁴Rehabilitation Teaching and Research Unit, University of Otago, Dunedin, New Zealand

⁵Institute of Applied Health Sciences, University of Aberdeen, Aberdeen, UK

⁶NHS Ayrshire and Arran, Kilmarnock, UK

⁷School of Health and Life Sciences, Glasgow Caledonian University, Glasgow, UK

⁸Health Services Research Unit, University of Aberdeen, Aberdeen, UK

⁹School of Nursing and Midwifery, Robert Gordon University, Aberdeen, UK

¹⁰Queen Elizabeth University Hospital, NHS Greater Glasgow and Clyde, Glasgow, UK

¹¹Centre for Healthcare Randomised Trials (CHaRT), Health Services Research Unit, University of Aberdeen, Aberdeen, UK

¹²Edinburgh Clinical Trials Unit, University of Edinburgh, Edinburgh, UK

*Corresponding author

Declared competing interests of authors: Sarah G Dean reports funding from University College London for consultancy work as part of an expert panel in the behaviour change technique taxonomy development, Medical Research Council-funded project (2012–2015) outside the submitted work. Mohamed Abdel-Fattah reports having been, in the past, a speaker, consultant and surgical trainer for Astellas Pharma Inc. (Tokyo, Japan), Bard Medical Division (Covington, GA, USA), Pfizer Inc. (New York, NY, USA), American Medical Systems Inc. (Minnetonka, MN, USA) and Coloplast Ltd (Peterborough, UK). At that time, he was reimbursed travel expenses, received limited personal honorariums for some work, and, on occasions, received sponsorship towards attending scientific conferences. He has received a research grant from Coloplast, managed by the or

received assistance University of Aberdeen. A limited number of his trainees attended pharmaceutical-sponsored educational/leadership workshops and/towards presenting their research work in scientific conferences. He was also the chairperson of the Scottish Pelvic Floor Network, which, at the time (up to 2014), received sponsorship from various industrial companies to exhibit in annual meetings and surgical workshops. In addition, he receives travel sponsorship from numerous national and international conferences and non-profit organisations when invited as guest speaker and/or expert surgeon. Wael Agur reports non-financial support from Astellas Pharma; personal fees from Johnson & Johnson (New Brunswick, NJ, USA) – Ethicon Endoscopy; personal fees from Specialty European Pharma Ltd/Contura Ltd (London, UK); personal fees and non-financial support from the Chief Scientist Office (Edinburgh, UK), NHS Research Scotland and NHS Ayrshire & Arran; grants from the National Institute for Health Research (NIHR); grants from the University of Stirling; personal fees from London Medical Education Academy (London, UK); personal fees from the Central Legal Office, NHS Scotland; personal fees from various law firms acting on behalf of mesh manufacturers, patient claimants and defendant clinicians in Scotland, England, Wales, Northern Ireland, the Republic of Ireland, USA and Australia; personal fees from Oaklaw Consultancy Ltd (Kilmarnock, UK) (dealing with medicolegal consultancy through a third party – Medico-legal Administration Service). Karen L Guerrero reports personal speaker fees, outside the submitted work, and educational bursaries from Contura Ltd to attend continuing professional development meetings. She is also course director/faculty member for Glasgow Urogynaecology/NHS Education Scotland courses/training days and meetings and NHS/Deanery Educational courses, which have had industry sponsorship. In addition, she is Treasurer for the UK Continence Society, the British Society of Urogynaecology Educational Committee Chair elect, Secretary of the Scottish Pelvic Floor Network and sits on the Royal College of Obstetricians and Gynaecologists subspecialty training committee. All of these associations/charities have meetings supported by industry. Furthermore, she is co-grant holder for several NIHR grants. John Norrie reports grants from the University of Aberdeen, and the University of Edinburgh during the conduct of the study; and declares membership of the following NIHR boards: Cardiopulmonary Resuscitation Decision-making Committee (2016), Health Technology Assessment (HTA) Commissioning Board (2010–16), HTA Commissioning Sub-Board (Expression of Interest) (2016–19), HTA Funding Boards Policy Group (2016–19), HTA General Board (2016–19), HTA Post-Board funding teleconference (2016–19), NIHR Clinical Trials Unit Standing Advisory Committee (2018–present), NIHR HTA and Efficacy and Mechanism Evaluation Editorial Board (2014–19) and the Pre-exposure Prophylaxis Impact Review Panel (2017–present).

Published December 2020

DOI: 10.3310/hta24700

This report should be referenced as follows:

Hagen S, Bugge C, Dean SG, Elders A, Hay-Smith J, Kilonzo M, *et al.* Basic versus biofeedback-mediated intensive pelvic floor muscle training for women with urinary incontinence: the OPAL RCT. *Health Technol Assess* 2020;**24**(70).

Health Technology Assessment is indexed and abstracted in *Index Medicus*/MEDLINE, *Excerpta Medica*/EMBASE, *Science Citation Index Expanded* (SciSearch®) and *Current Contents*®/Clinical Medicine.

ISSN 1366-5278 (Print)

ISSN 2046-4924 (Online)

Impact factor: 3.370

Health Technology Assessment is indexed in MEDLINE, CINAHL, EMBASE, the Cochrane Library and Clarivate Analytics Science Citation Index.

This journal is a member of and subscribes to the principles of the Committee on Publication Ethics (COPE) (www.publicationethics.org/).

Editorial contact: journals.library@nihr.ac.uk

The full HTA archive is freely available to view online at www.journalslibrary.nihr.ac.uk/hta. Print-on-demand copies can be purchased from the report pages of the NIHR Journals Library website: www.journalslibrary.nihr.ac.uk

Criteria for inclusion in the *Health Technology Assessment* journal

Reports are published in *Health Technology Assessment* (HTA) if (1) they have resulted from work for the HTA programme, and (2) they are of a sufficiently high scientific quality as assessed by the reviewers and editors.

Reviews in *Health Technology Assessment* are termed 'systematic' when the account of the search appraisal and synthesis methods (to minimise biases and random errors) would, in theory, permit the replication of the review by others.

HTA programme

Health Technology Assessment (HTA) research is undertaken where some evidence already exists to show that a technology can be effective and this needs to be compared to the current standard intervention to see which works best. Research can evaluate any intervention used in the treatment, prevention or diagnosis of disease, provided the study outcomes lead to findings that have the potential to be of direct benefit to NHS patients. Technologies in this context mean any method used to promote health; prevent and treat disease; and improve rehabilitation or long-term care. They are not confined to new drugs and include any intervention used in the treatment, prevention or diagnosis of disease.

The journal is indexed in NHS Evidence via its abstracts included in MEDLINE and its Technology Assessment Reports inform National Institute for Health and Care Excellence (NICE) guidance. HTA research is also an important source of evidence for National Screening Committee (NSC) policy decisions.

This report

The research reported in this issue of the journal was funded by the HTA programme as project number 11/71/03. The contractual start date was in September 2013. The draft report began editorial review in December 2018 and was accepted for publication in May 2019. The authors have been wholly responsible for all data collection, analysis and interpretation, and for writing up their work. The HTA editors and publisher have tried to ensure the accuracy of the authors' report and would like to thank the reviewers for their constructive comments on the draft document. However, they do not accept liability for damages or losses arising from material published in this report.

This report presents independent research funded by the National Institute for Health Research (NIHR). The views and opinions expressed by authors in this publication are those of the authors and do not necessarily reflect those of the NHS, the NIHR, NETSCC, the HTA programme or the Department of Health and Social Care. If there are verbatim quotations included in this publication the views and opinions expressed by the interviewees are those of the interviewees and do not necessarily reflect those of the authors, those of the NHS, the NIHR, NETSCC, the HTA programme or the Department of Health and Social Care.

© Queen's Printer and Controller of HMSO 2020. This work was produced by Hagen *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Published by the NIHR Journals Library (www.journalslibrary.nihr.ac.uk), produced by Prepress Projects Ltd, Perth, Scotland (www.prepress-projects.co.uk).

NIHR Journals Library Editor-in-Chief

Professor Ken Stein Professor of Public Health, University of Exeter Medical School, UK

NIHR Journals Library Editors

Professor John Powell Chair of HTA and EME Editorial Board and Editor-in-Chief of HTA and EME journals. Consultant Clinical Adviser, National Institute for Health and Care Excellence (NICE), UK, and Professor of Digital Health Care, Nuffield Department of Primary Care Health Sciences, University of Oxford, UK

Professor Andrée Le May Chair of NIHR Journals Library Editorial Group (HS&DR, PGfAR, PHR journals) and Editor-in-Chief of HS&DR, PGfAR, PHR journals

Professor Matthias Beck Professor of Management, Cork University Business School, Department of Management and Marketing, University College Cork, Ireland

Dr Tessa Crilly Director, Crystal Blue Consulting Ltd, UK

Dr Eugenia Cronin Senior Scientific Advisor, Wessex Institute, UK

Dr Peter Davidson Consultant Advisor, Wessex Institute, University of Southampton, UK

Ms Tara Lamont Senior Scientific Adviser (Evidence Use), Wessex Institute, University of Southampton, UK

Dr Catriona McDaid Senior Research Fellow, York Trials Unit, Department of Health Sciences, University of York, UK

Professor William McGuire Professor of Child Health, Hull York Medical School, University of York, UK

Professor Geoffrey Meads Emeritus Professor of Wellbeing Research, University of Winchester, UK

Professor John Norrie Chair in Medical Statistics, University of Edinburgh, UK

Professor James Raftery Professor of Health Technology Assessment, Wessex Institute, Faculty of Medicine, University of Southampton, UK

Dr Rob Riemsma Reviews Manager, Kleijnen Systematic Reviews Ltd, UK

Professor Helen Roberts Professor of Child Health Research, UCL Great Ormond Street Institute of Child Health, UK

Professor Jonathan Ross Professor of Sexual Health and HIV, University Hospital Birmingham, UK

Professor Helen Snooks Professor of Health Services Research, Institute of Life Science, College of Medicine, Swansea University, UK

Professor Ken Stein Professor of Public Health, University of Exeter Medical School, UK

Professor Jim Thornton Professor of Obstetrics and Gynaecology, Faculty of Medicine and Health Sciences, University of Nottingham, UK

Professor Martin Underwood Warwick Clinical Trials Unit, Warwick Medical School, University of Warwick, UK

Please visit the website for a list of editors: www.journalslibrary.nihr.ac.uk/about/editors

Editorial contact: journals.library@nihr.ac.uk

Abstract

Basic versus biofeedback-mediated intensive pelvic floor muscle training for women with urinary incontinence: the OPAL RCT

Suzanne Hagen,^{1*} Carol Bugge,² Sarah G Dean,³ Andrew Elders,¹ Jean Hay-Smith,⁴ Mary Kilonzo,⁵ Doreen McClurg,¹ Mohamed Abdel-Fattah,⁵ Wael Agur,⁶ Federico Andreis,² Joanne Booth,⁷ Maria Dimitrova,⁵ Nicola Gillespie,¹ Cathryn Glazener,⁸ Aileen Grant,⁹ Karen L Guerrero,¹⁰ Lorna Henderson,¹¹ Marija Kovandzic,² Alison McDonald,¹¹ John Norrie,¹² Nicole Sergenson,¹ Susan Stratton,¹ Anne Taylor² and Louise R Williams¹

¹Nursing, Midwifery and Allied Health Professions Research Unit, Glasgow Caledonian University, Glasgow, UK

²Faculty of Health Sciences and Sport, University of Stirling, Stirling, UK

³University of Exeter Medical School, Exeter, UK

⁴Rehabilitation Teaching and Research Unit, University of Otago, Dunedin, New Zealand

⁵Institute of Applied Health Sciences, University of Aberdeen, Aberdeen, UK

⁶NHS Ayrshire and Arran, Kilmarnock, UK

⁷School of Health and Life Sciences, Glasgow Caledonian University, Glasgow, UK

⁸Health Services Research Unit, University of Aberdeen, Aberdeen, UK

⁹School of Nursing and Midwifery, Robert Gordon University, Aberdeen, UK

¹⁰Queen Elizabeth University Hospital, NHS Greater Glasgow and Clyde, Glasgow, UK

¹¹Centre for Healthcare Randomised Trials (CHaRT), Health Services Research Unit, University of Aberdeen, Aberdeen, UK

¹²Edinburgh Clinical Trials Unit, University of Edinburgh, Edinburgh, UK

*Corresponding author hagen@gcu.ac.uk

Background: Urinary incontinence affects one in three women worldwide. Pelvic floor muscle training is an effective treatment. Electromyography biofeedback (providing visual or auditory feedback of internal muscle movement) is an adjunct that may improve outcomes.

Objectives: To determine the clinical effectiveness and cost-effectiveness of biofeedback-mediated intensive pelvic floor muscle training (biofeedback pelvic floor muscle training) compared with basic pelvic floor muscle training for treating female stress urinary incontinence or mixed urinary incontinence.

Design: A multicentre, parallel-group randomised controlled trial of the clinical effectiveness and cost-effectiveness of biofeedback pelvic floor muscle training compared with basic pelvic floor muscle training, with a mixed-methods process evaluation and a longitudinal qualitative case study. Group allocation was by web-based application, with minimisation by urinary incontinence type, centre, age and baseline urinary incontinence severity. Participants, therapy providers and researchers were not blinded to group allocation. Six-month pelvic floor muscle assessments were conducted by a blinded assessor.

Setting: This trial was set in UK community and outpatient care settings.

Participants: Women aged ≥ 18 years, with new stress urinary incontinence or mixed urinary incontinence. The following women were excluded: those with urgency urinary incontinence alone, those who had received formal instruction in pelvic floor muscle training in the previous year, those unable to contract their pelvic floor muscles, those pregnant or < 6 months postnatal, those with prolapse greater than stage II, those currently having treatment for pelvic cancer, those with cognitive impairment affecting capacity to give informed consent, those with neurological disease, those with a known nickel allergy or sensitivity and those currently participating in other research relating to their urinary incontinence.

Interventions: Both groups were offered six appointments over 16 weeks to receive biofeedback pelvic floor muscle training or basic pelvic floor muscle training. Home biofeedback units were provided to the biofeedback pelvic floor muscle training group. Behaviour change techniques were built in to both interventions.

Main outcome measures: The primary outcome was urinary incontinence severity at 24 months (measured using the International Consultation on Incontinence Questionnaire Urinary Incontinence Short Form score, range 0–21, with a higher score indicating greater severity). The secondary outcomes were urinary incontinence cure/improvement, other urinary and pelvic floor symptoms, urinary incontinence-specific quality of life, self-efficacy for pelvic floor muscle training, global impression of improvement in urinary incontinence, adherence to the exercise, uptake of other urinary incontinence treatment and pelvic floor muscle function. The primary health economic outcome was incremental cost per quality-adjusted-life-year gained at 24 months.

Results: A total of 300 participants were randomised per group. The primary analysis included 225 and 235 participants (biofeedback and basic pelvic floor muscle training, respectively). The mean 24-month International Consultation on Incontinence Questionnaire Urinary Incontinence Short Form score was 8.2 (standard deviation 5.1) for biofeedback pelvic floor muscle training and 8.5 (standard deviation 4.9) for basic pelvic floor muscle training (adjusted mean difference -0.09 , 95% confidence interval -0.92 to 0.75 ; $p = 0.84$). A total of 48 participants had a non-serious adverse event (34 in the biofeedback pelvic floor muscle training group and 14 in the basic pelvic floor muscle training group), of whom 23 (21 in the biofeedback pelvic floor muscle training group and 2 in the basic pelvic floor muscle training group) had an event related/possibly related to the interventions. In addition, there were eight serious adverse events (six in the biofeedback pelvic floor muscle training group and two in the basic pelvic floor muscle training group), all unrelated to the interventions. At 24 months, biofeedback pelvic floor muscle training was not significantly more expensive than basic pelvic floor muscle training, but neither was it associated with significantly more quality-adjusted life-years. The probability that biofeedback pelvic floor muscle training would be cost-effective was 48% at a £20,000 willingness to pay for a quality-adjusted life-year threshold. The process evaluation confirmed that the biofeedback pelvic floor muscle training group received an intensified intervention and both groups received basic pelvic floor muscle training core components. Women were positive about both interventions, adherence to both interventions was similar and both interventions were facilitated by desire to improve their urinary incontinence and hindered by lack of time.

Limitations: Women unable to contract their muscles were excluded, as biofeedback is recommended for these women.

Conclusions: There was no evidence of a difference between biofeedback pelvic floor muscle training and basic pelvic floor muscle training.

Future work: Research should investigate other ways to intensify pelvic floor muscle training to improve continence outcomes.

Trial registration: Current Controlled Trial ISRCTN57746448.

Funding: This project was funded by the NIHR Health Technology Assessment programme and will be published in full in *Health Technology Assessment*; Vol. 24, No. 70. See the NIHR Journals Library website for further project information.

Contents

List of tables	xiii
List of figures	xv
List of abbreviations	xvii
Plain English summary	xix
Scientific summary	xxi
Chapter 1 Introduction	1
Scientific background and current evidence base	1
<i>Possible mechanisms for how biofeedback might improve continence</i>	1
<i>Evidence for the effectiveness of biofeedback for female urinary incontinence</i>	2
<i>Rationale for research</i>	2
<i>Aims and objectives</i>	2
<i>Development of the OPAL trial interventions</i>	3
Chapter 2 Trial design and methods	5
Trial design	5
Ethics approval and research governance	5
Management of the trial	5
Participants and setting	5
Inclusion criteria	5
Exclusion criteria	6
Recruitment procedure	6
Informed consent	6
Randomisation, concealment and blinding	6
Therapist training	7
Treatment group allocation	7
<i>Both trial groups (biofeedback PFMT and basic PFMT)</i>	7
<i>Biofeedback PFMT</i>	7
Data collection and management	8
Participant follow-up	8
Outcome measures	9
<i>Primary outcome</i>	9
<i>Secondary outcomes</i>	9
<i>Economic-related outcomes</i>	10
Adverse events	10
Sample size	11
Statistical methods	11
Important changes to the project protocol	12
Chapter 3 Trial outcomes and results	13
Participants	13
Baseline characteristics	14
Intervention received	14
Primary outcome measure	18

Sensitivity analyses	18
Subgroup analyses	20
Urinary outcomes	20
Further treatment for urinary incontinence	23
Pelvic floor outcomes	24
Adherence	26
Adverse events	27
Chapter 4 Health economic evaluation	29
Methods	29
<i>Data collection</i>	29
<i>Intervention costs</i>	30
<i>Medication costs</i>	30
<i>Unit costs</i>	30
<i>Estimation of participant costs</i>	31
<i>Participant- and companion-incurred costs and indirect costs</i>	31
<i>Quality-of-life measures</i>	32
<i>Data analysis</i>	33
<i>Missing data</i>	33
<i>Incremental cost per quality-adjusted life-year gained</i>	33
<i>Sensitivity and subgroup analyses</i>	34
Results	34
<i>Missing data</i>	35
<i>Resource use</i>	35
<i>Intervention</i>	35
<i>Follow-up care</i>	36
<i>Intervention and follow-up costs</i>	37
<i>Generic and condition-specific quality-of-life measures</i>	37
<i>Costs incurred by and quality-adjusted life-years of cured and incontinent participants</i>	38
<i>Costs directly incurred by participants and carers and indirect costs of time and productivity lost due to urinary incontinence</i>	40
<i>Base-case cost-effectiveness analysis</i>	40
<i>Sensitivity analysis</i>	42
<i>Subgroup analysis</i>	43
<i>Type of incontinence</i>	45
<i>Age</i>	46
<i>Severity of incontinence</i>	46
<i>Quality-of-life scores</i>	46
<i>Discounting and unadjusted costs and quality-adjusted life-years</i>	46
<i>Assuming that data were missing not at random</i>	46
Discussion	46
Conclusion	48
Chapter 5 Process evaluation	49
Introduction	49
Methods	49
Process evaluation sampling, recruitment strategy and data collection	49
<i>Appointment checklists</i>	49
<i>Appointment audio-recordings</i>	50
<i>Therapist interviews</i>	50
<i>PFMT exercise diaries</i>	50

Process evaluation analysis	51
<i>Appointment checklists</i>	51
<i>Appointment audio-recordings</i>	51
<i>Therapist interviews</i>	51
<i>PFMT exercise diaries</i>	51
Results	51
<i>Appointment checklists</i>	51
<i>Appointment audio-recordings</i>	52
<i>Therapist interviews</i>	55
<i>Participant diaries</i>	60
Discussion	61
Conclusion	62
Chapter 6 Longitudinal qualitative case study	65
Introduction	65
Methods	65
<i>Sampling and recruitment</i>	65
<i>Case study data collection</i>	66
<i>Case study data analysis</i>	66
<i>Management and governance</i>	67
Findings	67
<i>Sample</i>	67
<i>Women's adherence to the interventions</i>	68
<i>Facilitators of adherence during the active treatment phase</i>	68
<i>Barriers to intervention adherence during active treatment phase</i>	71
<i>Facilitators of women's adherence in the maintenance phase</i>	73
<i>Barriers to women's adherence in the maintenance phase</i>	75
<i>Women's urinary incontinence outcomes in the short and long term</i>	77
<i>Theoretical propositions</i>	78
Discussion	78
<i>Strengths and limitations of the case study</i>	78
<i>Comparison of findings to existing literature</i>	79
Chapter 7 Results synthesis	81
Aims	81
Methods	81
Results	81
Discussion	84
Chapter 8 Discussion and conclusions	85
Statement of principal findings	85
<i>Robustness of primary outcome analysis</i>	85
<i>Further analyses of the primary outcome</i>	85
Strength and limitations	86
<i>Strengths</i>	86
<i>Limitations</i>	86
Interpretation of results	87
Implications for health care	88
Future research implications	89
Acknowledgements	91
References	95

Appendix 1 Trial components	101
Appendix 2 Overview of participant data collected via questionnaires	103
Appendix 3 Recruitment by centre	105
Appendix 4 Reasons for ineligibility	107
Appendix 5 Sensitivity analysis under assumptions of missing not at random	109
Appendix 6 Subgroup analysis	111
Appendix 7 Patient Global Impression of Improvement in urinary incontinence	113
Appendix 8 Bowel and prolapse symptoms	115
Appendix 9 Estimation of unit cost of biofeedback machines per participant	119
Appendix 10 Unit costs of patient-reported medications	121
Appendix 11 Unit cost of incontinence products (pads that women wear and disposable bed pads)	123
Appendix 12 Generalised linear model diagnostics test results	125
Appendix 13 Missing data (costs)	127
Appendix 14 Missing data (utility measures)	129
Appendix 15 Proportion of participants experiencing any problems on each of the EQ-5D-3L domains: analysis based on all of the available EQ-5D-3L data points	131
Appendix 16 Proportion of participants who reported issues on each domain of the ICIQ-LUTSqol: analysis based on all of the available EQ-5D-3L data points	133
Appendix 17 Sensitivity analysis cost-effectiveness acceptability curves and incremental cost-effectiveness ratio scatterplots	135
Appendix 18 Examples of core and optional components of the basic PFMT and biofeedback PFMT intervention protocols, as worded on the therapist assessment form checklists	143

List of tables

TABLE 1 Baseline characteristics	16
TABLE 2 Urinary measures at baseline	16
TABLE 3 Pelvic floor measures at baseline	17
TABLE 4 Appointment attendance and therapist type	18
TABLE 5 Home exercise programme recommended	19
TABLE 6 Summary of ICIQ-UI SF responses and analysis of differences between groups (models correct for minimisation variables and baseline ICIQ-UI SF)	19
TABLE 7 Lower urinary tract symptoms (ICIQ-FLUTS and ICIQ-LUTSqol)	21
TABLE 8 Cure and improvement in urinary symptoms	22
TABLE 9 Further treatment for UI	23
TABLE 10 Hospital admissions and non-surgical treatment for UI	24
TABLE 11 Self-efficacy for PFMT	25
TABLE 12 Pelvic floor muscle function assessed by therapist via vaginal examination at baseline and 6 months (blinded)	25
TABLE 13 Therapists' adherence to intervention delivery protocol at appointments and women's adherence to prescribed intervention at home	27
TABLE 14 Participant-reported frequency of PFMEs	28
TABLE 15 Average NHS unit costs	30
TABLE 16 Participant time and travel unit cost	32
TABLE 17 Missing resource use/cost and QALY data	35
TABLE 18 Summary of intervention and follow-up care resource use (complete-case data)	36
TABLE 19 Summary of intervention and follow-up costs (complete-case data)	37
TABLE 20 Mean EQ-5D-3L scores at each time point, QALYs and differences between treatment groups	38
TABLE 21 Mean ICIQ-LUTSqol scores at each time point, second-year QALYs and differences between treatment groups	40
TABLE 22 Direct and indirect costs to participants and carers (complete-case data)	42

TABLE 23 Base-case cost-effectiveness results: therapist-reported duration of appointment (imputed data set)	42
TABLE 24 Recommended appointment time (3.5 hours for six appointments) (imputed data set)	43
TABLE 25 Sensitivity analyses	44
TABLE 26 Number of available checklists compared with potential total number of checklists by group and by appointment	52
TABLE 27 Summary of checklist core components ticked for each appointment by group: differentiating between basic PFMT core components and biofeedback-specific PFMT core components	53
TABLE 28 Summary of checklist optional components ticked for each appointment by group: differentiating between basic PFMT optional components and biofeedback-specific PFMT optional components	54
TABLE 29 Summary of audio-recordings made by appointment and by group	55
TABLE 30 Summary of audio-recorded evidence of delivery of core components by group and by appointment	56
TABLE 31 Characteristics of women in the case study by group allocation	67
TABLE 32 Case study examples of variation in adherence to treatment by allocated treatment group and across time	68
TABLE 33 Case study examples of variance in UI outcomes at 24 months by allocated treatment group	77
TABLE 34 Convergence coding matrix	82
TABLE 35 Bowel symptoms	115
TABLE 36 Prolapse symptoms (POP-SS)	118

List of figures

FIGURE 1 Projected and actual recruitment	13
FIGURE 2 The CONSORT flow diagram	15
FIGURE 3 Sensitivity analyses	20
FIGURE 4 Subgroup analyses	21
FIGURE 5 Proportion of participants experiencing any problems in each of the EQ-5D-3L domains	39
FIGURE 6 Proportion of participants experiencing any problems in each of the ICIQ-LUTSqol domains by group	41
FIGURE 7 Cost-effectiveness acceptability curves: therapist-reported duration of appointment (imputed data set)	42
FIGURE 8 Scatterplot of incremental costs and QALYs for biofeedback PFMT vs. basic PFMT: therapist-reported duration of appointment (imputed data set)	43
FIGURE 9 Cost-effectiveness acceptability curves: recommended appointment time (3.5 hours for six appointments)	43
FIGURE 10 Scatterplot of incremental costs and QALYs for biofeedback PFMT vs. basic PFMT: recommended appointment time (3.5 hours for six appointments)	44
FIGURE 11 Proportion of exercise diaries signed (by participant and therapist) and returned by appointment and by group (in respect to the total number attending that appointment in that group)	60
FIGURE 12 Pattern mixture model under assumptions of MNAR	109
FIGURE 13 Cost-effectiveness acceptability curve: SUI subgroup analysis	135
FIGURE 14 Scatterplot of incremental cost and incremental QALYs: SUI subgroup analysis	135
FIGURE 15 Cost-effectiveness acceptability curve: MUI subgroup analysis	136
FIGURE 16 Scatterplot of incremental costs and incremental QALYs: MUI subgroup analysis	136
FIGURE 17 Cost-effectiveness acceptability curve: age < 50 years subgroup analysis	136
FIGURE 18 Scatterplot of incremental costs and QALYs: age < 50 years subgroup analysis	137
FIGURE 19 Cost-effectiveness acceptability curve: age \geq 50 years subgroup analysis	137

FIGURE 20 Scatterplot of incremental costs and incremental QALYs: age ≥ 50 years subgroup analysis	137
FIGURE 21 Cost-effectiveness acceptability curve: lower severity incontinence level (ICIQ-UI SF score of < 13) subgroup analysis	138
FIGURE 22 Scatterplot of incremental cost and incremental QALYs: lower severity incontinence level (ICIQ-UI SF score of < 13) subgroup analysis	138
FIGURE 23 Cost-effectiveness acceptability curve: higher severity incontinence level (ICIQ-UI SF score of ≥ 13) subgroup analysis	138
FIGURE 24 Scatterplot of incremental costs and incremental QALYs: higher severity incontinence level (ICIQ-UI SF score of ≥ 13) subgroup analysis	139
FIGURE 25 Cost-effectiveness acceptability curve: using ICIQ-LUTSqol scores as a utility measure	139
FIGURE 26 Scatterplot of incremental costs and incremental QALYs: using ICIQ-LUTSqol scores as a utility measure	139
FIGURE 27 Cost-effectiveness acceptability curve: undiscounted costs and effects	140
FIGURE 28 Scatterplot of incremental costs and incremental QALYs: undiscounted costs and effects	140
FIGURE 29 Cost-effectiveness acceptability curve: discounted costs and effects at 6%	140
FIGURE 30 Scatterplot of incremental costs and incremental QALYs: discounted costs and effects at 6%	141
FIGURE 31 Cost-effectiveness acceptability curve: unadjusted cost-effectiveness analysis	141
FIGURE 32 Scatterplot of incremental cost and incremental QALYs: unadjusted cost-effectiveness analysis	141

List of abbreviations

AE	adverse event	ITT	intention to treat
BCT	behaviour change technique	MNAR	missing not at random
CACE	complier-average causal effect	MUI	mixed urinary incontinence
CEAC	cost-effectiveness acceptability curve	NICE	National Institute for Health and Care Excellence
CHaRT	Centre for Healthcare Randomised Trials	NIHR	National Institute for Health Research
CI	confidence interval	NMAHP RU	Nursing, Midwifery and Allied Health Professions Research Unit
CONSORT	Consolidated Standards of Reporting Trials	NMB	net monetary benefit
DMEC	Data Monitoring and Ethics Committee	OPAL	Optimal Pelvic floor muscle training for Adherence Long-term
EMG	electromyography	OR	odds ratio
EQ-5D-3L	EuroQol-5 Dimensions, three-level version	PFME	pelvic floor muscle exercise
GCU	Glasgow Caledonian University	PFMT	pelvic floor muscle training
GLM	generalised linear model	PGI-I	Patient Global Impression of Improvement
GP	general practitioner	PGI-S	Patient Global Impression of Severity
HCHS	Hospital and Community Health Service	PhD	Doctor of Philosophy
ICER	incremental cost-effectiveness ratio	PMG	Project Management Group
ICIQ-FLUTS	International Consultation on Incontinence Questionnaire for Female Lower Urinary Tract Symptoms	POP-SS	pelvic organ prolapse symptom score
ICIQ-LUTSqol	International Consultation on Incontinence Questionnaire for Lower Urinary Tract Symptoms Quality of Life	QALY	quality-adjusted life-year
ICIQ-UI SF	International Consultation on Incontinence Questionnaire Urinary Incontinence Short Form	R&D	research and development
IMB	information motivation behavioural skills	RCT	randomised controlled trial
IT	information technology	SAE	serious adverse event
		SD	standard deviation
		SUI	stress urinary incontinence
		TAF	therapist assessment form
		TSC	Trial Steering Committee
		UI	urinary incontinence
		WTP	willingness to pay

Plain English summary

Urinary incontinence (accidental leakage of urine) is a common and embarrassing problem for women. Pregnancy and childbirth may contribute by leading to less muscle support and bladder control. Pelvic floor exercises and 'biofeedback' equipment (a device that lets women see the muscles working as they exercise) are often used in treatment. There is good evidence that exercises (for the pelvic floor) can help, but less evidence about whether or not adding biofeedback provides better results.

This trial compared pelvic floor exercises alone with pelvic floor exercises plus biofeedback. Six hundred women with urinary incontinence participated. Three hundred women were randomly assigned to the exercise group and 300 women were randomised to the exercise plus biofeedback group. Each woman had an equal chance of being in either group. Women were offered six appointments with a therapist over 16 weeks to receive their allocated treatment.

After 2 years, there was no difference between the groups in the severity of women's urinary incontinence. Women in both groups varied in how much exercise they managed to do. Some managed to exercise consistently over the 2 years and others less so. There were many factors (other than the treatment received) that affected a woman's ability to exercise. Notably, women viewed the therapists' input very positively. The therapists reported some problems fitting biofeedback into the appointments, but, overall, they delivered both treatments as intended. Women carried out exercises at home and many in the biofeedback pelvic floor muscle training group also used biofeedback at home; however, for both groups, time issues, forgetting and other health problems affected their adherence. There were no serious complications related to either treatment.

Overall, exercise plus biofeedback was not significantly more expensive than exercise alone and the quality of life associated with exercise plus biofeedback was not better than the quality of life for exercise alone.

In summary, exercises plus biofeedback was no better than exercise alone. The findings do not support using biofeedback routinely as part of pelvic floor exercise treatment for women with urinary incontinence.

Scientific summary

Background

Urinary incontinence (accidental urine leakage) is a distressing problem that affects around one in three women. The main types of urinary incontinence are stress (involuntary urine leakage associated with exertion or effort, or with sneezing or coughing), urgency (involuntary urine leakage accompanied by, or immediately preceding, a compelling desire to pass urine) and mixed (combined stress and urgency), with stress being the most common. Current UK guidelines recommend that women with urinary incontinence are offered at least 3 months of pelvic floor muscle training. There is evidence that pelvic floor muscle training is effective in treating urinary incontinence; however, it is not clear how intensively women have to exercise their pelvic floor muscles to give the maximum sustained improvement in symptoms, and how to enable women to achieve this. Electromyography biofeedback is an adjunct to pelvic floor muscle training that may help women exercise more intensively for longer and, thus, may improve continence outcomes when compared with pelvic floor muscle training alone.

The need for a definitive trial of biofeedback as an intensifier of pelvic floor muscle training is supported by the current evidence base. The Cochrane systematic review of biofeedback-assisted pelvic floor muscle training for women with urinary incontinence (Herderschee R, Hay-Smith EJ, Herbison GP, Roovers JP, Heineman MJ. Feedback or biofeedback to augment pelvic floor muscle training for urinary incontinence in women. *Cochrane Database Syst Rev* 2011;**7**:CD009252) found that biofeedback-assisted pelvic floor muscle training appeared to offer benefit over basic pelvic floor muscle training; however, this effect may have been confounded by the greater amount of health professional contact in the biofeedback groups within trials. Thus, a trial in which women in both groups have similar amounts of contact is needed to allow firm conclusions to be drawn about biofeedback as an intensifier of pelvic floor muscle training in its own right.

The Optimal Pelvic floor muscle training for Adherence Long-term (OPAL) trial was designed to address this gap by answering a pragmatic question concerned with the clinical effectiveness and cost-effectiveness of electromyography biofeedback-mediated intensive pelvic floor muscle training (biofeedback pelvic floor muscle training) compared with pelvic floor muscle training alone (basic pelvic floor muscle training) for the treatment of stress or mixed female urinary incontinence.

Objective

The overall objective of the trial was to evaluate the clinical effectiveness and cost-effectiveness of electromyography biofeedback as an adjunct to pelvic floor muscle training, the latter being the recommended first-line treatment for female urinary incontinence in the UK. We conducted a randomised controlled trial in which electromyography biofeedback was delivered in addition to pelvic floor muscle training as an intensifier (both in clinic and at home), and compared with pelvic floor muscle training alone. We included a process evaluation to identify mediating factors, such as intervention fidelity, that might affect the clinical effectiveness of the intervention, and to establish how these factors influence clinical effectiveness and whether or not they differ between randomised groups. A nested longitudinal qualitative case study was also included to explore women's experiences of the trial interventions, to identify the barriers to and facilitators of adherence, to establish how they influence adherence and to assess whether or not they differ between randomised groups.

Methods

We carried out an individually randomised, multicentre controlled trial in which we compared biofeedback pelvic floor muscle training with basic pelvic floor muscle training for women with stress or mixed urinary incontinence to assess superiority. The allocation ratio was 1 : 1. The allocation was carried out remotely via a web-based automated application, with minimisation by type of urinary incontinence, centre, age and urinary incontinence severity at baseline. A sample size of 600 women was needed to detect a difference of 2.5 points on the primary outcome measure between the groups, with 90% power and 5% significance level, assuming a standard deviation of 10 and allowing for attrition of > 20%.

Participants were recruited from UK centres in community and outpatient care settings, where continence care is usually provided. During the trial recruitment period, women attending for their first continence appointment were identified by the health-care team at each centre. Eligibility was assessed by a clinician who saw the woman at a screening appointment, which included a vaginal examination. The inclusion criteria were being female, being aged ≥ 18 years and newly presenting with stress or mixed urinary incontinence. Women excluded were those with urgency urinary incontinence alone, those who had received formal instruction in pelvic floor muscle training in the previous year, those unable to contract their pelvic floor muscles, those who were pregnant or < 6 months postnatal, those with prolapse greater than stage II, those currently having treatment for pelvic cancer, those with cognitive impairment affecting capacity to give informed consent, those with neurological disease, those with a known nickel allergy or sensitivity and those currently participating in other research relating to their urinary incontinence.

The primary outcome measure was severity of urinary incontinence at 24 months post randomisation, as measured by the validated International Consultation on Incontinence Questionnaire – Urinary Incontinence Short Form. Interim follow-ups were conducted at 6 and 12 months. Secondary outcome measures included number of participants who were cured/improved, uptake of other treatment for urinary incontinence, presence of other pelvic floor symptoms (bladder, bowel, prolapse), self-efficacy for and adherence to pelvic floor muscle training, and quality of life. Resource use data were collected by participant self-report at the follow-up time points. The primary health economic outcome measure of cost-effectiveness was incremental cost per quality-adjusted life-year at 24 months.

Participants were offered six appointments with a trained therapist over 16 weeks, during which the trial intervention (biofeedback pelvic floor muscle training or basic pelvic floor muscle training) was delivered. Biofeedback units were provided to participants in the biofeedback pelvic floor muscle training group for home use. A written intervention manual was prepared for the therapists to guide intervention delivery. A checklist was provided for each appointment, detailing the intervention components to be delivered. Behaviour change techniques, both core and optional, were built in to the intervention protocols delivered (e.g. to help participants with goal-setting, action-planning and problem-solving).

As part of the process evaluation, therapists delivering the intervention at each centre were interviewed. In addition, to record their delivery of the intervention, therapists completed an intervention checklist at every appointment with a participant. A selection of appointments from across centres, where therapists were delivering the trial interventions, were audio-recorded. A subset of participants took part in a longitudinal case study and were interviewed at baseline and at 6, 12 and 24 months. These interviews were transcribed and analysed at different levels: within time points, within-case, cross-case and between trial groups. Quantitative descriptive summaries and framework analysis methods were used to analyse these mixed-methods data sets.

Results

Twenty-three centres agreed to take part in the trial and 687 women were screened for eligibility: a total of 600 women were randomised (300 women to the biofeedback pelvic floor muscle training

group and 300 women to the basic pelvic floor muscle training group). The two randomised groups were comparable at baseline. Follow-up questionnaire return rates were 74% at 6 months, 84% at 12 months and 78% at 24 months. Adherence to the intervention appointments was good: 92% attended at least one appointment (95% biofeedback pelvic floor muscle training, 89% basic pelvic floor muscle training) and 36% attended all six appointments (37% biofeedback pelvic floor muscle training, 36% basic pelvic floor muscle training). The average number of appointments attended was 4.1 (4.2 biofeedback pelvic floor muscle training, 4.0 basic pelvic floor muscle training).

The primary analysis indicated that there was no difference between groups in the severity of urinary incontinence at 24 months (mean difference -0.09 , 95% confidence interval -0.92 to 0.75 ; $p = 0.84$). This finding was robust to sensitivity analysis that investigated assumptions about non-compliance and missing data. There were no significant differences between groups in the primary outcome within the predefined subgroups of type of urinary incontinence, age, baseline severity or type of therapist. There were no significant differences between groups in International Consultation on Incontinence Questionnaire Urinary Incontinence Short Form score at 6 or 12 months.

The analysis of secondary outcomes indicated no significant differences between groups, at any time point, in urinary incontinence cure or improvement, other lower urinary tract symptoms, condition-specific quality of life, participant impression of improvement, uptake of further treatment for urinary incontinence, pelvic floor function, prolapse symptoms or bowel symptoms. The biofeedback pelvic floor muscle training group had statistically significantly better scores for self-efficacy for pelvic floor muscle exercises at 24 months, but the difference was small and unlikely to be clinically significant.

There were 48 participants who had a non-serious adverse event (34 biofeedback pelvic floor muscle training, 14 basic pelvic floor muscle training), of whom 23 (21 biofeedback pelvic floor muscle training, 2 basic pelvic floor muscle training) had a complication related or possibly related to one of the interventions. There were eight serious adverse events (6 biofeedback pelvic floor muscle training, 2 basic pelvic floor muscle training), all of which were unrelated.

The base-case economic analysis concluded that, at 24 months, biofeedback pelvic floor muscle training (£956) was not significantly more expensive than basic pelvic floor muscle training (£906) (mean difference £50, 95% confidence interval $-£84$ to $£184$), but neither was it associated with significantly more quality-adjusted life-years (1.567 vs. 1.566, mean difference -0.0009 , 95% confidence interval -0.06 to 0.06). The incremental cost-effectiveness ratio of £56,617 per quality-adjusted life-year gained exceeded society's willingness-to-pay threshold of £30,000. Biofeedback pelvic floor muscle training was associated with a 49% chance of being cost-effective if society was willing to pay £30,000 for a quality-adjusted life-year.

The process evaluation had two key findings. First, therapists did deliver a more intensive intervention to the biofeedback pelvic floor muscle training group, despite time pressures. Second, most participants in both groups did receive core behavioural change techniques embedded in the basic pelvic floor muscle training, as intended, but few optional behavioural change techniques were used by therapists in either group.

Women from both groups who were interviewed reported positive experiences of the interventions received. There was variation in adherence and urinary incontinence outcomes in both groups, with no clear differences between groups. Factors that helped women adhere in the short and long term that were similar in both groups were their desire to improve their continence, having the belief that they were able to undertake pelvic floor muscle training and the expectation of what they might achieve in terms of symptoms and quality of life. Adherence was facilitated by the therapists and their input was particularly valued. Having time to undertake pelvic floor muscle training and 'life taking over' were important barriers to short- and long-term adherence.

Conclusions

Implications for health care

- We can be confident in concluding that the addition of biofeedback to pelvic floor muscle training treatment does not improve incontinence severity at 2 years and is unlikely to be cost-effective. Therefore, we think it unlikely that routinely offering this adjunct to pelvic floor muscle training will benefit continence outcomes for women.
- It is feasible for the behaviour change techniques that were embedded in the basic pelvic floor muscle training (which have evidence of effectiveness in other fields) to be used within the context of pelvic floor muscle training for women with urinary incontinence.

Recommendations for research (in priority order)

- Investigate other intensive forms of pelvic floor muscle training to improve continence outcomes (e.g. a pelvic floor muscle training programme with the addition of more health professional support).
- Establish if there is potential for evaluation of newer biofeedback devices in the context of treatment of urinary incontinence.

Trial registration

This trial is registered as ISRCTN57746448.

Funding

Funding for this study was provided by the Health Technology Assessment programme of the National Institute for Health Research.

Chapter 1 Introduction

Scientific background and current evidence base

There are various options for the management of female urinary incontinence (UI), including behavioural approaches, pelvic floor muscle training (PFMT), medication, surgery, nerve stimulation, injectable bulking agents, a botulinum toxin injection to the bladder wall, vaginal continence pessaries and the use of containment products, such as pads and catheters. These options may be used in combination.

Currently in the UK, supervised PFMT of at least 3 months' duration is the first-line treatment for stress urinary incontinence (SUI) and mixed urinary incontinence (MUI).¹ 'Supervised PFMT' refers to a course of treatment that involves one-to-one supervision by a health-care professional (usually a physiotherapist or nurse) who carries out a digital vaginal examination to ensure that the woman can correctly contract her pelvic floor muscles, and provides ongoing teaching, guidance and motivation relating to pelvic floor muscle exercises (PFMEs). The potential annual spend on PFMT is estimated at £38M (0.03% of the NHS budget of £106B), based on 0.8% of women aged ≥ 15 years being referred to UI services each year.² PFMT refers to the regular practice of repeated and progressive pelvic floor muscle contractions in order to produce a training effect on the muscles. The aim of a PFMT programme is to increase functionality of the pelvic floor through several mechanisms, detailed by Bo *et al.*³

Adjuncts used with supervised PFMT in clinical practice include vaginal cones, electrical stimulation and biofeedback. Biofeedback is the technique by which information about a normally unconscious physiological process is presented to the patient and/or the therapist as a visual, auditory or tactile signal.⁴ Electromyography (EMG) is the study of minute electrical potentials produced by depolarisation of muscle membrane.⁵ In EMG biofeedback, which is the focus of this research, electrical activity arising from muscle activity (during exercise and voluntary effort) is recorded in microvolts and displayed as a visual and/or auditory signal for both the patient and therapist to view. In a PFMT programme, an internal vaginal or anal probe is used to record electrical information from pelvic floor muscles through surface recordings. The probe is connected by cables to a biofeedback unit. Handheld units, with a small visual display screen, are available for home use. The display provides a visual representation of the muscles contracting and relaxing, allowing monitoring of strength, endurance and repetitions.

Possible mechanisms for how biofeedback might improve continence

Biofeedback acts as a strong and specific training stimulus, which has the potential to markedly intensify a training programme. This effect is evidenced in the biofeedback and general exercise literature.⁶ In this report, the exercise being referred to is PFME. In the context of supervised PFMT for treatment of UI, biofeedback contributes to:

- teaching women the precise, correct muscle contraction in terms of technique, force and timing, and assisting the therapist in prescribing the correct exercise dose to ensure that appropriate physiological responses occur in the muscles
- informing adjustments to the exercise prescribed by the therapist (modulating)
- highlighting to women improvements in their muscle function (encouraging)⁷
- increasing the quality of the interaction with the therapist (via biofeedback used in ways congruent with health behaviour change theory and practice)
- change in women's behaviour relating to home exercise, as biofeedback has been classified as a behaviour change technique (BCT) in a taxonomy of BCTs.⁸

These features have the potential to increase a woman's confidence (self-efficacy) in doing PFMT; self-efficacy for PFMT is a determinant of intention to adhere⁹ and adherence itself.^{10,11} For PFMT to have an effect on UI, sufficient exercise must be carried out over a long enough period to strengthen (hypertrophy) the muscle.¹² Thus, maximising adherence to exercise is key to an intensive PFMT intervention. Intensive PFMT, which fosters PFMT self-efficacy, can potentially be achieved through the addition of biofeedback, leading to improvement in the quality and quantity of PFMT performed by women, and to more improvement in the severity of their UI than a less intensive programme.

Evidence for the effectiveness of biofeedback for female urinary incontinence

Both a Cochrane systematic review of biofeedback-assisted PFMT for women with UI⁷ and a Health Technology Assessment-funded systematic review of non-surgical treatments for women with SUI¹³ found that biofeedback-assisted PFMT appeared to offer benefit over basic PFMT alone. However, the Cochrane review,⁷ with its more restricted scope, conducted a more detailed analysis of the biofeedback trials and found that this effect may be confounded by the greater amount of health professional contact received by women in the biofeedback groups.

Rationale for research

Based on current evidence, it is not clear if the apparent benefit of biofeedback can be attributed to the biofeedback or to some other variable, such as additional health professional contact. Some of the trials in the Cochrane review⁷ had the same PFMT programme in both arms of the trial and the same amount of health professional contact. The results from these trials still favoured biofeedback, although the difference between groups was not statistically significant. A common problem in all these trials was the failure to clearly state the purpose of biofeedback, or to describe the intervention protocol. Thus, it was not clear if biofeedback could, theoretically or in practice, change the clinical effectiveness of the PFMT. Currently, the routine use of biofeedback as part of PFMT is not recommended and should be considered only for women who are unable to contract their pelvic floor muscles.¹ Despite this, we know that therapists use biofeedback in their clinical practice as an adjunct to a supervised PFMT programme when treating female UI.¹⁴

Thus, a robust comparison of biofeedback-mediated intensive PFMT compared with basic PFMT, in which both groups have the same basic PFMT programme and the same amount of health professional contact, is imperative to establish whether or not biofeedback does add value and improve incontinence outcomes.

Aims and objectives

The overall aim of the Optimal Pelvic floor muscle training for Adherence Long-term (OPAL) trial was to carry out a randomised controlled trial (RCT) to evaluate a theoretically based biofeedback-intensified PFMT intervention to answer the primary research question: is biofeedback-mediated intensive PFMT, compared with basic PFMT, a clinically effective and cost-effective treatment for women with SUI or MUI (stress and urgency)?

The trial comprised the following components:

- Development of biofeedback-mediated intensive PFMT and basic PFMT interventions based on health behaviour theory designed to treat SUI and MUI.
- A RCT in which the use of biofeedback as an adjunct to basic PFMT (biofeedback PFMT) was compared with PFMT alone (basic PFMT), in terms of clinical effectiveness and cost-effectiveness.
- A process evaluation to identify and investigate the possible mediating factors that affect the effectiveness of the interventions (including fidelity to intervention delivery and uptake), how these mediating factors influence effectiveness and whether or not the factors differ between randomised groups.
- A longitudinal qualitative case study to investigate women's experiences of the trial interventions, to identify the barriers and facilitators that affect adherence, to explain the process through which they influence adherence and to identify whether or not these differ between randomised groups.

Development of the OPAL trial interventions

The OPAL trial interventions (including biofeedback-mediated intensive PFMT and basic PFMT protocols) were developed by Jean Hay-Smith, Doctor of Philosophy (PhD) (University of Otago, Dunedin, New Zealand); Sarah G Dean, PhD (University of Exeter, Exeter, UK); Doreen McClurg, PhD [Nursing, Midwifery and Allied Health Professions Research Unit (NMAHP RU) Research Unit, Glasgow Caledonian University (GCU), Glasgow, UK]; Suzanne Hagen, PhD (NMAHP RU, GCU, Glasgow UK); Carol Bugge, PhD (University of Stirling, Stirling, UK); and Joanne Booth, PhD (GCU, Glasgow, UK).

The OPAL trial interventions are essentially behavioural intervention (i.e. an intervention focusing on the adoption and maintenance of a PFMT programme, with or without use of biofeedback). PFMT adherence is problematic and often decreases over time.¹⁵ PFMT interventions typically include a number of 'active' ingredients that 'interact' (e.g. in the OPAL trial, the exercise and biofeedback ingredients are presumed to interact to intensify PFMT) to produce an effect that is potentially greater than the sum of the parts. Thus, the OPAL interventions had two core characteristics (behavioural difficulty and interacting components) of a 'complex' intervention.¹⁶ A further element of complexity was the tailoring components within the interventions, which allowed flexibility in delivery; these components have been called 'optional' in the intervention protocols.

Four sources were used to develop the intervention:

1. theory consistent with our objective of changing health behaviour: the information motivation behavioural skills (IMB) model¹⁷
2. existing qualitative evidence about women's experiences of PFMT and why they did or did not adopt the behaviour to undertake PFMT¹⁸
3. the explicit use of BCTs to support adoption and maintenance of the behaviour⁸
4. the existing PFMT intervention from the Pelvic Organ Prolapse Physiotherapy (POPPY) trial,¹⁹ which provided the outline for what therapists considered 'basic' PFMT in UK clinical practice.

Once developed, the intervention protocols to be delivered were detailed in the intervention manual, which was given to each therapist. Separate protocols were available for biofeedback-mediated intensive PFMT (referred to hereafter as biofeedback PFMT) and basic PFMT. In practice, the interventions were delivered over six appointments with a trial therapist, which the participant attended. Details of the desired content of each appointment were provided in the intervention manual and in checklist format in the therapist assessment form (TAF), completed by the therapist at each appointment. This is described in more detail in *Chapter 3*.

Chapter 2 Trial design and methods

Trial design

The OPAL trial was designed to compare biofeedback PFMT with basic PFMT, in terms of long-term continence severity [see the project web page: www.journalslibrary.nihr.ac.uk/programmes/hta/117103/#/ (accessed 29 July 2019)]. It included a superiority multicentre RCT comparing two treatment groups: (1) PFMT with EMG biofeedback in clinic and at home (biofeedback PFMT) and (2) PFMT alone (basic PFMT)²⁰ (see *Chapter 3*).

The main trial was supplemented with a separate economic evaluation (see *Chapter 4*), a process evaluation (see *Chapter 5*) and a longitudinal qualitative case study (see *Chapter 6*). For a trial overview, see *Appendix 1*. A published protocol is available for the process evaluation and case study.²¹

Ethics approval and research governance

Ethics approval for all aspects of the trial was granted by the West of Scotland Research Ethics Committee 4 on 13 March 2013 [reference number 13/WS/0048; see the project web page: www.journalslibrary.nihr.ac.uk/programmes/hta/117103/#/ (accessed 29 July 2019)]. Overall research and development (R&D) approval for the trial from NHS Research Scotland was granted on 24 January 2014 (reference number NRS13/UR13) and from the National Institute for Health Research (NIHR) Co-ordinated System for gaining NHS permission on 2 April 2014 (reference number 120377). Local NHS R&D approval for the trial was granted by 16 different trusts in England, and six local health boards granted R&D approval for seven centres in Scotland. The trial sponsor was GCU and the OPAL trial office was based in the NMAHP RU at GCU. The OPAL trial is registered with the International Standard Randomised Controlled Trial Register (ISRCTN57746448).

Management of the trial

A Project Management Group (PMG), made up of all co-applicants and research staff employed on the trial, met regularly, face to face or by teleconference, to review the trial's progress. An independent Trial Steering Committee (TSC) and Data Monitoring and Ethics Committee (DMEC) met at least yearly to review trial progress. Patient and public involvement in the trial was ensured by involvement in the development and delivery of the trial, and membership of the PMG, of a lay person who had experience of treatment for female UI. A second lay person was a member of the TSC.

Participants and setting

The trial recruited women with SUI or MUI from 23 centres across the UK, comprising hospital outpatient and community settings.

Inclusion criteria

The trial included women:

- aged ≥ 18 years, presenting with a new episode of SUI or MUI.

Exclusion criteria

The trial excluded women:

- with urgency UI alone
- who had received formal instruction in PFMT in the previous year
- not able to contract their pelvic floor muscles
- who were pregnant or < 6 months postnatal
- with a prolapse greater than stage II (i.e. > 1 cm below the hymen on valsalva)
- receiving active treatment for pelvic cancer
- with cognitive impairment affecting capacity to give informed consent
- who had the following neurological diseases: multiple sclerosis, Parkinson's disease, stroke, motor neurone disease, spinal cord injury
- with a known nickel allergy or sensitivity
- who were already participating in other research relating to UI.

Recruitment procedure

The research team at each centre was responsible for identifying potential participants from women newly presenting with UI, with urinary leakage as the presenting complaint. A pre-screening process was used to exclude women who were obviously not eligible and therefore not approached to be formally screened at a clinic appointment. Screening could take place either at the first clinical presentation or a separate appointment arranged to assess eligibility. During screening appointments, clinicians assessed women for a clinical diagnosis of SUI or MUI and performed a vaginal examination to assess other trial eligibility criteria. With ethics approval, these details, along with contact details, were recorded on the clinical assessment form [see the project web page: www.journalslibrary.nihr.ac.uk/programmes/hta/117103/#/ (accessed 29 July 2019)]. Each potentially eligible woman was allocated a unique trial identity number that, if the woman was recruited, was used throughout the trial in all the woman's trial materials. Centre trial staff discussed the trial with women who were eligible and willing to take part, provided a patient information leaflet [see the project web page: www.journalslibrary.nihr.ac.uk/programmes/hta/117103/#/ (accessed 29 July 2019)], answered any questions and provided women with a consent form to complete [see the project web page: www.journalslibrary.nihr.ac.uk/programmes/hta/117103/#/ (accessed 29 July 2019)].

Informed consent

Informed, written consent was sought after the women had received all trial information and after any questions about participating had been answered. Copies of the completed consent form were held (1) by the participating woman, (2) in their general practitioner's (GP's) notes, (3) at the centre and (4) at the trial office. Women were provided with information about the process evaluation audio-recordings and the linked case study [see the project web page: www.journalslibrary.nihr.ac.uk/programmes/hta/117103/#/ (accessed 29 July 2019)] and given the option of being approached to take part in these additional aspects of the research (see *Chapters 5 and 6*).

Randomisation, concealment and blinding

Consenting participants were individually randomised to one of two groups: biofeedback PFMT or basic PFMT. Group allocation was generated by the web-based randomisation service provided by the Centre for Healthcare Randomised Trials (CHaRT), a UK Clinical Research Collaboration clinical trials unit, located in the Health Services Research Unit, University of Aberdeen, Aberdeen, UK. A woman's group allocation was relayed by e-mail to the trial office and recruiting member of staff at the centre. Owing to the nature

of the intervention, blinded allocation was not possible for the participants and therapists delivering the intervention. The data entry staff and statistician were not blinded, as some data collection forms were intentionally different between groups, and this would be apparent to those staff processing the data. However, clinicians performing the 6-month pelvic floor muscle assessments were blinded to group allocation. Group allocation used minimisation based on four variables: UI type (SUI or MUI), recruiting centre, age ($< 50/\geq 50$ years) and UI severity [International Consultation on Incontinence Questionnaire Urinary Incontinence Short Form (ICIQ-UI SF) score of < 13 or ≥ 13].

Therapist training

To ensure consistency of expertise, all therapists received face-to-face training in delivery of both interventions (biofeedback PFMT and basic PFMT, as described in *Treatment group allocation*) by OPAL trial staff. They were also supplied with the OPAL trial intervention manual, providing a background to the OPAL trial, and the OPAL intervention protocols, describing the health behaviour model adopted to maximise intervention adherence.

Treatment group allocation

Both trial groups (biofeedback PFMT and basic PFMT)

In both groups, participants completed the basic PFMT protocol over a 16-week period:²² six individual face-to-face appointments with a therapist were offered at weeks 0, 1, 3, 6, 10 and 15, based on a previous trial of PFMT.¹⁹ At the first appointment the therapist performed a vaginal examination, including visual inspection and digital assessment of the pelvic floor muscles (the Oxford Classification,²³ and the International Continence Society method²⁴).

During this assessment, all participants were taught how to contract and relax their pelvic floor muscles and how to pre-contract their pelvic floor muscles prior to (and in expectation of) increases in abdominal pressure, such as coughing and sneezing ('the Knack'²⁵).

Based on the findings of the assessment, the therapist and participant agreed a PFMT programme, tailored according to the woman's ability and lifestyle, which progressed over time. The PFMT programme focused on strength and endurance, aiming for three sets of contractions per day, progressed by increasing the number of repetitions, hold duration and body position modification on progression (i.e. lying, sitting, standing, squatting).

As per local practice, therapists provided participants with information to support good bladder management and methods of dealing with urgency and frequency. Therapists also used the BCTs that had been mapped to, and detailed in, the intervention protocols to encourage participants' adherence to the PFMT programme. Participants recorded the PFMT they performed in a diary [see the project web page: www.journalslibrary.nihr.ac.uk/programmes/hta/117103/#/ (accessed 29 July 2019)].

No EMG biofeedback equipment was used during appointments for the basic PFMT group. Verbal feedback based on digital vaginal palpation was, however, permitted.

Biofeedback PFMT

In addition to the basic PFMT protocol provided to both groups [see *Both trial groups (biofeedback PFMT and basic PFMT)*], women randomised to the biofeedback PFMT group received a biofeedback protocol during appointments and were provided with a biofeedback unit to use at home. Biofeedback units were provided by the trial office and the same type of unit was used during appointments and at home

[NeuroTrac® Simplex single-channel EMG biofeedback unit (Verity Medical Ltd, Romsey, UK)]. The biofeedback unit displayed participants' pelvic floor muscle activity, providing audio and visual feedback on muscle contraction and relaxation. Participants were instructed how to insert, use and clean the probe, operate the unit and interpret the biofeedback information. Therapists set the parameters of the biofeedback unit and agreed with participants the frequency for use at home. The units stored data on home usage and participants also recorded their home use (and PFMT practice) in a diary.

Data collection and management

Baseline characteristics were recorded for each participant, including age, body mass index, number of births and delivery mode (number of breech, caesarean, forceps, normal vaginal and vacuum deliveries) and type of UI and severity (normal, mild, moderate, severe). Participants also recorded their urinary leakage at baseline in a 3-day bladder diary [see the project web page: www.journalslibrary.nihr.ac.uk/programmes/hta/117103/#/ (accessed 29 July 2019)].

Baseline and 6-, 12- and 24-month data were collected via questionnaire booklets completed by the participants, containing the primary and secondary outcome measures [see the project web page: www.journalslibrary.nihr.ac.uk/programmes/hta/117103/#/ (accessed 29 July 2019)]. See *Appendix 2* for an overview of the time points of when participant data were collected via questionnaires.

At each appointment, participants were given a diary to record the PFMT carried out (and biofeedback use, where appropriate), during the intervention period. Participants had the telephone number of their therapist written on their diary so they could make contact if there were any concerns about the home exercise programme.

At a participant's first appointment, therapists recorded demographic information, medical history and pelvic floor assessment findings in the TAF. At each subsequent appointment the therapist used the TAF to record pelvic floor assessment findings, the treatment plan, the prescribed PFMT programme and the participant's adherence.

Participants attended an appointment at 6 months for a further pelvic floor assessment. This was undertaken by an assessor who was blinded to the participant's group allocation and had not been involved in treatment delivery. Blinding was achieved by instructing participants in their appointment letter not to discuss their allocated treatment group with the assessor until after the assessment was complete. Assessors were directed in a trial standard operating procedure to remind participants not to discuss their trial treatment until after the examination. Assessors also recorded whether or not they had been aware of a participant's group when carrying out the examination and, if so, to provide details.

Researchers at the trial office entered the data described above into the trial database, developed by CHaRT.

Participant follow-up

The duration of follow-up was 24 months from date of randomisation. It is important to assess duration of effect; previous trials typically measured outcomes immediately post treatment or followed up participants for a maximum of only 1 year. Follow-up questionnaires were sent from the trial office by post or e-mail to participants, with postal, e-mail and telephone reminders. During telephone reminder calls, participants could verbally complete a shortened version of follow-up questionnaires, including ICIQ-UI SF, EuroQol-5 Dimensions, three-level version (EQ-5D-3L), and questions about PFMT adherence and UI treatment uptake [see the project web page: www.journalslibrary.nihr.ac.uk/programmes/hta/117103/#/ (accessed 29 July 2019)].

Outcome measures

Primary outcome

The primary outcome was the ICIQ-UI SF score at 24 months. This met with the commissioning brief from the funder, which highlighted symptom severity as an important outcome and 2 years as the minimum duration of follow-up. The ICIQ-UI SF²⁶ is a four-item questionnaire (total score ranging from 0 to 21, with higher scores indicating greater severity), covering frequency of UI (never = 0, once a week = 1, two or three times a week = 2, once a day = 3, several times a day = 4, all the time = 5), amount of leakage (none = 0, small = 2, moderate = 4, large = 6), overall impact of UI (not at all = 0, a great deal = 10) and a diagnostic question to identify type of UI. Responses to the first three questions are summed to give the total score. The maximum score is 21; the higher the score, the more severe the UI, with a score of ≥ 13 rated as severe.²⁷

Secondary outcomes

Urinary outcomes

The number of women with UI cured and improved was derived from the ICIQ-UI SF (cured = negative response to both 'how often do you leak urine?' and 'how much urine do you usually leak'; improved = a reduction in ICIQ-UI SF score of ≥ 3 points).²⁸ The Patient Global Impression of Improvement (PGI-I), which has been validated for use in females with UI,²⁹ was completed to assess perceptions of improvement in UI. The PGI-I enabled participants to describe their UI at each time point compared with the start of the trial, using a 7-point scale from 1 ('very much better') to 7 ('very much worse'). The International Consultation on Incontinence Questionnaire for Female Lower Urinary Tract Symptoms (ICIQ-FLUTS)³⁰ was completed to capture other urinary symptoms (nocturia, urgency, frequency, SUI, unexplained UI, bladder pain, diurnal frequency, hesitancy, straining, intermittency and enuresis), comprising 12 items, each scored 0–4, and three subscales [filling score (range 0–16), voiding score (range 0–12) and incontinence score (range 0–20)].

Participants completed questions assessing uptake of other treatment and use of services for UI: hospital admission, outpatient appointment, GP consultation, nurse appointment, physiotherapy appointment, surgery, medication and other treatment/medical advice.

Quality-of-life outcomes

Condition-specific quality of life was measured using the International Consultation on Incontinence Questionnaire for Lower Urinary Tract Symptoms Quality of Life (ICIQ-LUTSqol),³¹ which has 19 items, each scored from 1 to 4 (range 19–76). General health was measured using the EQ-5D-3L questionnaire³² (range –0.594 to 1) and the EuroQol-5 Dimensions (EQ-5D) visual analogue score (range 0–100).

Pelvic floor-related outcomes

Pelvic floor outcomes were assessed using four different measurements: (1) the pelvic organ prolapse symptom score (POP-SS), a seven-item questionnaire measuring prolapse symptoms;³³ (2) bowel symptoms were assessed via an unpublished early version of the International Consultation on Incontinence Questionnaire Bowel Short Form, comprising six items covering frequency and consistency of bowel movement; (3) the Oxford Classification²³ was used by therapists to quantify pelvic floor muscle strength on a 6-point scale, as well as contraction endurance and repetitions; and (4) therapists also used the International Continence Society classification for muscle relaxation (absent, partial, complete) and contraction (absent, weak, normal, strong).²⁴

Self-efficacy for PFMT

Women's self-efficacy for PFMT was assessed using the PFME self-efficacy scale,³⁴ a 17-item self-report scale comprising two factors: (1) belief in PFME execution and its benefits (11 items); and (2) belief in performing PFME as scheduled and despite barriers (six items). Items are scored from 1 (strongly disagree) to 5 (strongly agree). Higher scores are associated with a woman perceiving that she has greater confidence in her ability to perform PFMEs.

Adherence to PFMT

Adherence to performing PFMT and the use of biofeedback was captured through the number of appointments attended, participants completing PFMT diaries during the 16-week intervention period, information in the TAFs on whether or not the recommended intervention protocol had been followed and participants' responses to questions in the follow-up questionnaires. Additional understanding of perceived outcomes, self-efficacy and adherence was gained through the process evaluation (see *Chapter 5*) and case study (see *Chapter 6*).

Economic-related outcomes

The primary health economic outcome measure of cost-effectiveness was incremental cost per quality-adjusted life-year (QALY) at 24 months, based on responses to the EQ-5D-3L. This incremental cost-effectiveness ratio (ICER) was defined by the difference in cost between two trial interventions, divided by the difference in their effect.

Secondary outcomes were (1) the cost of resources used by participants, including the use of primary (GP services) and secondary (outpatient visits, inpatient stay, surgical interventions for incontinence) care services, and further referral for subsequent additional specialist management; (2) the personal costs to the participants, including costs of travelling to appointments and work/social restrictions; and (3) QALYs derived using the ICIQ-LUTSqol responses.

Full cost-effectiveness methods are described in detail in *Chapter 4*.

Adverse events

The interventions delivered in the trial are well established clinically; therefore, any adverse events (AEs) may commonly be found in women receiving PFMT and biofeedback. Expected AEs arising from the interventions were:

- sore pelvic floor muscles
- lower back pain
- vaginal irritation or discomfort
- thrush
- urinary tract infection
- spotting or staining (not linked to menstruation), potentially caused by biofeedback probe insertion
- vaginal itchiness and discomfort (potentially linked to biofeedback probe if nickel sensitivity/allergy)
- psychological distress from vaginal examination and/or use of biofeedback probe (e.g. as a result of previous abuse or distressing labour).

All AEs (for which the participant sought health-care professional interventions) and serious adverse events (SAEs) (an untoward medical occurrence in a participant, including death, life-threatening conditions, hospitalisation or prolonging of existing hospitalisation, persistent or significant disability or incapacity, a congenital anomaly or birth defect, and any event considered medically significant by the investigator) were assessed to determine cause, severity and relatedness, and were reported to the relevant regulatory bodies.

Serious adverse events were reported to the main Research Ethics Committee if they occurred within 30 days of the woman's previous therapy appointment and were, in the opinion of the chief investigator or the chairperson of the DMEC, related to trial participation (resulted from the administration of any of the research procedures) and unexpected (not listed in the protocol as an expected occurrence). In addition, SAE forms were used to record deaths from any cause during the course of the trial.

Sample size

There were no published long-term outcome data for our primary outcome measure to inform our sample size calculation; thus, we referred to studies reporting baseline ICIQ-UI SF data for women with SUI/MUI.^{35,36} These studies indicated a standard deviation (SD) of 5, but we expected that the SD at the 24-month time point in the trial could possibly be as high as 10. A minimal clinically important difference of 2.5 points on the ICIQ-UI SF score was assumed (a change in frequency of urine leakage of, e.g., 'once a day' to 'never'), based on a study of older women available at the time.³⁷ On this basis, a sample size of 234 participants per group would detect this difference, with 90% power at a significance level of 0.05. Allowing for a 22% dropout, a target of 300 women per group was set.

Statistical methods

The analysis was based on an intention-to-treat (ITT) principle, in which participants were analysed according to their randomised group, regardless of the intervention received. All outcomes were described with the appropriate descriptive statistics: means and SDs for continuous outcomes, counts and percentages for dichotomous and categorical outcomes.

The analysis of the primary outcome estimated the mean difference [with 95% confidence intervals (CIs)] in the ICIQ-UI SF score at 24 months between the biofeedback PFMT group and the basic PFMT group, using a linear mixed model, adjusting for the minimisation covariates of therapist type (physiotherapist or other type of therapist) and baseline score. Recruiting centre was fitted as a random effect. Assumptions of linearity and normality of error distributions were examined by inspection of residual plots. Statistical significance was at the 5% level. Equivalent analyses were conducted for the ICIQ-UI SF score at 6 and 12 months.

The effect of missing ICIQ-UI SF data was investigated under various assumptions in a set of sensitivity analyses. Multiple imputation and repeated-measures models were analysed under a missing at random assumption. Pattern mixture models were then used under a missing not at random (MNAR) assumption. In these analyses, the imputed ICIQ-UI SF scores were first adjusted by adding 2.5 points (the minimal clinically important difference) and then by subtracting 2.5 points. These adjustments were then repeated in only one group and repeated again by applying the adjustments in only the other group. A further sensitivity analysis of the primary outcome was conducted, which took non-compliers into account, using a complier-average causal effect (CACE) model.

Secondary outcomes were analysed in a similar manner to the analysis of the ICIQ-UI SF, using appropriate generalised linear models (GLMs) (linear mixed models for continuous outcomes, binary logistic regression for dichotomous outcomes and ordinal logistic regression for ordered categorical outcomes). When ordinal models were fitted, the proportional odds assumption was examined using a Brant test. All models adjusted for minimisation variables, therapist type and, if measured, baseline score.

Subgroup analyses of the primary outcome by type of incontinence (SUI or MUI), type of therapist, participant age (< 50/≥ 50 years) and UI severity (ICIQ-UI SF score of < 13 or ≥ 13) were also conducted. A stricter level of statistical significance (1%) was set for the subgroup analyses, reflecting their exploratory nature. Heterogeneity of treatment effects among subgroups was tested for, using the appropriate subgroup by treatment group interactions.

The analysis of the trial data were conducted according to the statistical analysis plan [see the project web page: www.journalslibrary.nihr.ac.uk/programmes/hta/117103/#/ (accessed 29 July 2019)], which was written by investigators blinded to accruing outcome data and agreed by the DMEC and the TSC. The DMEC reviewed confidential reports of accumulating data and a single main analysis was performed at the end of the trial when 24-month follow-up was complete and the database locked. We did not plan

or conduct any interim analyses of outcome data, based on a decision made at the beginning of the trial and following discussion with the DMEC. Deviations from the statistical analysis plan are documented [see the project web page: www.journalslibrary.nihr.ac.uk/programmes/hta/117103/#/ (accessed 29 July 2019)].

Important changes to the project protocol

The original protocol included the following exclusion criterion: women who were pregnant or were < 1 year postnatal. This was subsequently relaxed to women who were < 6 months postnatal. This recognised that any natural resolution of postnatal UI would occur within the first 6 months and, therefore, beyond this time point women should be eligible for inclusion. The original criteria: women who have had formal instruction in PFMT in the previous 3 years was changed to in the previous year. Memory of PFMT instruction given more than 1 year ago was thought likely to be poor; therefore, these women would be similar to women who had not received instruction before. Women taking antimuscarinic medication were originally excluded; however, this did not reflect current clinical practice, so inclusion criteria were changed to enable women using this medication to take part. Furthermore, an additional criterion was agreed, excluding women with a known nickel allergy or sensitivity, as such women, if randomised to the biofeedback PFMT group, might experience irritation due to the biofeedback probe. A nickel-free probe could have been sourced for these women, but they were not routinely stocked by the NHS and we did not procure them because of the cost.

The original protocol specified that participants would be asked to complete a 3-day bladder diary at 24 months to quantify urine leakage. This aspect of data collection was stopped because of a poor response rate and the observation that it was impeding responses to the 24-month questionnaire (which was sent to participants at the same time) and having a negative impact on the primary outcome data completeness.

All changes to the protocol were approved by the TSC and DMEC, when appropriate, and the Research Ethics Committee.

Chapter 3 Trial outcomes and results

This chapter reports the results of the trial analysis set out in the statistical analysis plan [see the project web page: www.journalslibrary.nihr.ac.uk/programmes/hta/117103/#/ (accessed 29 July 2019)].

Participants

The OPAL trial recruited to target, with 600 participants randomised (300 to each group) between February 2014 and July 2016 (*Figure 1*). The trial had planned to recruit to target by November 2015, but the recruitment phase needed to be extended in order to randomise the required number of participants. Participants were recruited at 23 centres across the UK (see *Appendix 3*).

A total of 687 women were formally screened for eligibility, of whom 87 were ineligible or declined to participate (see *Appendix 4* for details of ineligibility). Of the eligible 629 women, 29 consented to participate but did not attend the subsequent appointment at which they would have been randomised. These women were therefore excluded from the trial, leaving 600 women who were randomised.

All outcome data were returned and entered by 4 June 2018, the date of agreed database lock.

Questionnaires were completed by 589 participants (98%) at baseline (292 in the biofeedback PFMT group and 297 in the basic PFMT group). Seven out of 600 participants withdrew consent to their data being used after randomisation, leaving 295 participants in the biofeedback PFMT group and 298 participants in the basic PFMT group included in the analysis (although the data for participants who withdrew are included in the summary of the baseline characteristics). At 6 months, 444 participants (74%) provided follow-up questionnaire data, rising to 504 participants (84%) at 12 months. At 24 months, the target of 468 responses

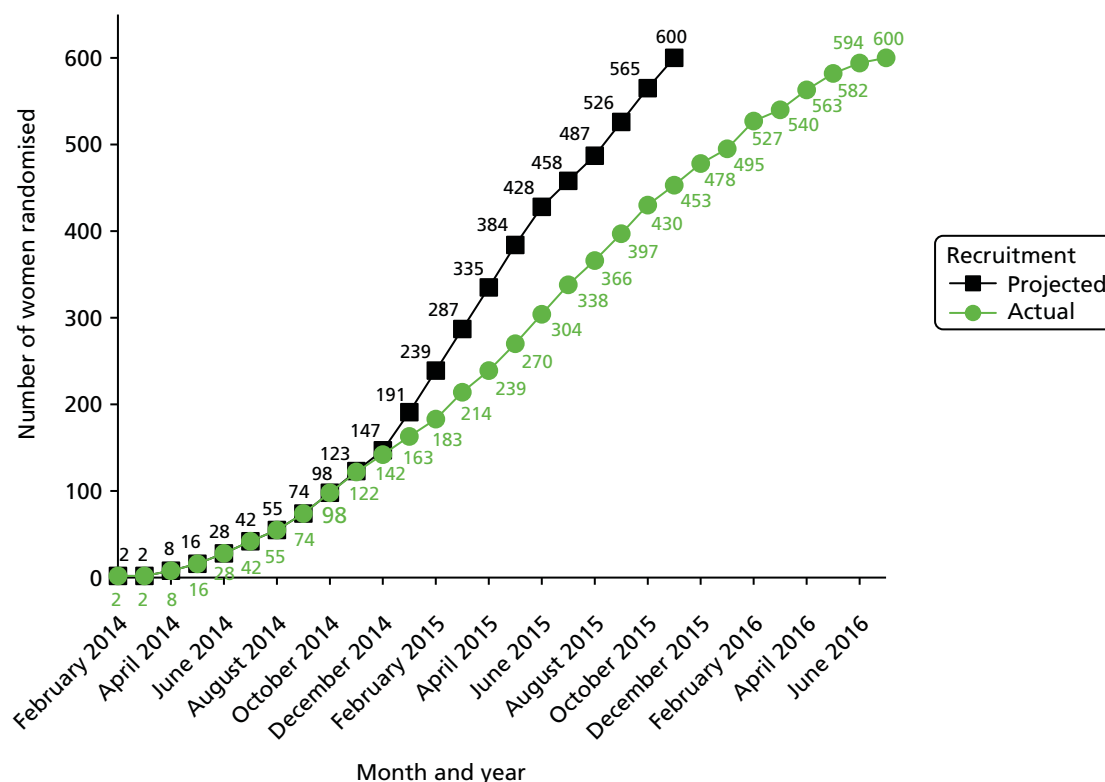


FIGURE 1 Projected and actual recruitment.

(78%) set out in the sample size calculation was achieved, with 230 (77%) in the biofeedback PFMT group and 238 (79%) in the basic PFMT group. Of those women who were followed up, the proportions of participants who completed the full questionnaire (rather than a shortened version over the telephone) were 82%, 75% and 72% at 6, 12 and 24 months, respectively. The response rates were similar between the two groups at each follow-up time point, for both methods of response. No participants died during the course of the trial. The number of participants at each stage of the trial is summarised in the Consolidated Standards of Reporting Trials (CONSORT) flow diagram (*Figure 2*).

Baseline characteristics

Table 1 summarises the baseline characteristics of participants. The age of participants ranged from 20 to 83 years, with a mean age of 47 years (48.2 ± 11.6 years in the biofeedback PFMT group and 47.3 ± 11.4 years basic PFMT group); 61.3% of participants in each group had MUI (SUI and urgency UI).

Baseline UI measures are reported in *Table 2*. The ICIQ-UI SF score (the primary outcome measure) ranged from 0 to the maximum of 21. Approximately half of the participants (51%) had UI at baseline that could be classed as severe (ICIQ-UI SF score of ≥ 13).²⁷ For the Patient Global Impression of Severity (PGI-S) scale, 55% of participants reported a rating of moderate or severe. Baseline pelvic floor measures show that < 8% of participants had an Oxford Scale score of ≥ 4 for slow contraction strength (*Table 3*). The Oxford Scale is measured from 0 to 5, with higher scores indicating stronger muscle function.²³ The mean self-efficacy for PFMT score was 62 (measured on a scale from 17 to 85, with higher scores indicating stronger beliefs in ability to exercise)³⁴ and 48% reported doing pelvic floor exercises at least one a week during the month prior to their participation in the trial.

Intervention received

The proportion of participants attending the maximum number of six appointments was 36.9% in the biofeedback PFMT group and 35.6% in the basic PFMT group (*Table 4*). The mean number of appointments was 4.2 ± 1.9 (biofeedback PFMT) and 4.0 ± 2.1 (basic PFMT). Forty-nine out of 593 women (8.3%) did not attend any appointments [$n = 16$ (5.4%) biofeedback PFMT; $n = 33$ (11.1%) basic PFMT].

The majority of participants (84.8%) were treated by only physiotherapists (86.8% biofeedback PFMT, 82.9% basic PFMT), with the remainder mostly being treated by only nurses. In most cases, the participant was treated by the same therapist throughout. There were 13 participants whose therapist type is classified as 'other and mixture', which included participants treated either by a midwife, a consultant or a mixture of different therapist types. All therapists participating in the trial were female.

According to our a priori definition, participants were classed as receiving 'treatment as allocated' (i.e. compliant with protocol or 'on-treatment') if pelvic floor muscle contractions were taught and feedback was given (at the first appointment), and a recommended exercise programme was written in the home exercise diary and given to the participant (during at least one appointment). In addition, participants in the biofeedback PFMT group were required to be taught insertion and removal of the probe and placement of the electrode at the first appointment in order to be classed as having treatment as allocated. This protocol compliance rate was 74.4% (198/266) in the biofeedback PFMT group and 83.3% (230/276) in the basic PFMT group. In a post hoc analysis, the definition for the biofeedback PFMT group was relaxed to allow for training in the use of the biofeedback device during either the first or second appointment, which reflected more accurately the instructions provided to therapists. Mostly, this was because it was not always possible to fit everything into the first appointment, but sometimes it was a clinical judgement that it was not appropriate for the participant; thus, therapists were told that biofeedback could be initiated in the second appointment. This revised definition resulted in a protocol compliance rate of 83.5% (222/266).

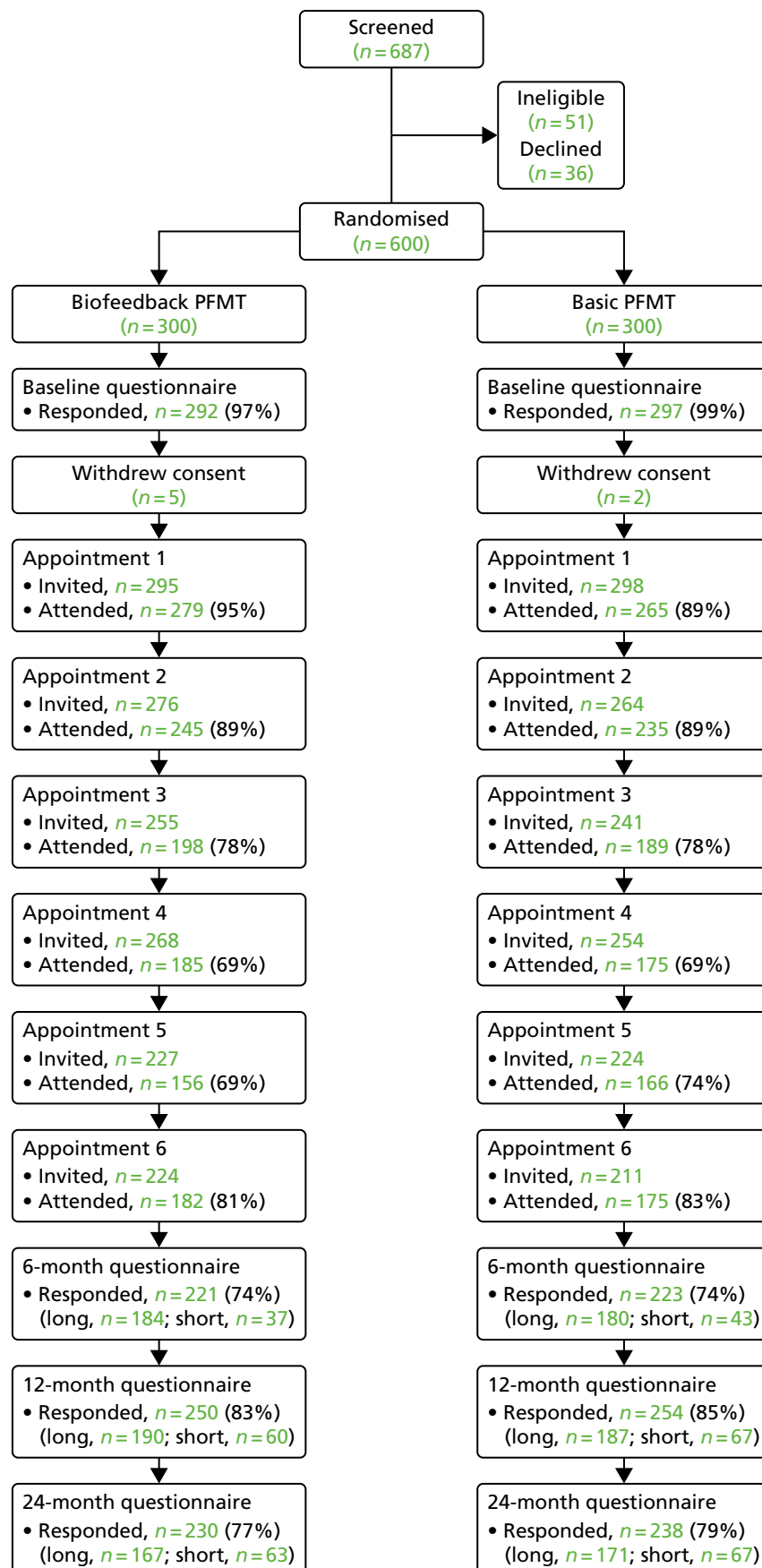


FIGURE 2 The CONSORT flow diagram.

TABLE 1 Baseline characteristics

Variable	Treatment group	
	Biofeedback PFMT	Basic PFMT
Age (years), <i>n</i> , mean (SD)	300, 48.2 (11.6)	300, 47.3 (11.4)
BMI (kg/m ²), <i>n</i> , mean (SD)	290, 28.6 (5.9)	287, 28.3 (6.2)
Number of births, <i>n/N</i> (%)		
0	21/298 (7.0)	12/289 (4.2)
1	40/298 (13.4)	60/289 (20.8)
2	116/298 (38.9)	122/289 (42.2)
3	83/298 (27.9)	63/289 (21.8)
≥ 4	38/298 (12.8)	32/289 (11.1)
Delivery mode history, <i>n/N</i> (%)		
Vaginal deliveries only	163/277 (58.8)	164/266 (61.7)
Caesarean deliveries only	11/277 (4.0)	15/266 (5.6)
Vaginal and caesarean deliveries	31/277 (11.2)	17/266 (6.4)
Any forceps delivery	52/277 (18.8)	55/266 (20.7)
Any vacuum delivery (but no forceps)	20/277 (7.2)	15/266 (5.6)
Type of incontinence, <i>n/N</i> (%)		
SUI	116/300 (38.7)	116/300 (38.7)
MUI	184/300 (61.3)	184/300 (61.3)
BMI, body mass index.		

TABLE 2 Urinary measures at baseline

Variable	Treatment group	
	Biofeedback PFMT	Basic PFMT
ICIQ-UI SF, <i>n</i> , mean (SD)	291, 12.5 (4.1)	294, 12.3 (3.7)
ICIQ-UI SF severity, <i>n/N</i> (%)		
Mild/moderate (score of < 13)	140/291 (48.1)	149/294 (50.7)
Severe (score of ≥ 13)	151/291 (51.9)	145/294 (49.3)
ICIQ-FLUTS filling score, <i>n</i> , mean (SD)	289, 5.0 (2.8)	297, 4.8 (2.6)
ICIQ-FLUTS voiding score, <i>n</i> , mean (SD)	292, 2.0 (2.0)	294, 2.0 (2.1)
ICIQ-FLUTS incontinence score, <i>n</i> , mean (SD)	290, 9.8 (3.6)	294, 9.3 (3.4)
PGI-S scale, <i>n/N</i> (%)		
Normal	13/292 (4.5)	23/294 (7.8)
Mild	113/292 (38.7)	115/294 (39.1)
Moderate	137/292 (46.9)	133/294 (45.2)
Severe	29/292 (9.9)	23/294 (7.8)
ICIQ-LUTSqol, <i>n</i> , mean (SD)	292, 43.5 (12.3)	297, 42.3 (12.1)
ICIQ-LUTSqol bother scale, <i>n</i> , mean (SD)	288, 7.4 (2.6)	288, 7.6 (2.5)
PGI-S, Patient Global Impression of Severity. All the urinary outcomes in this table are measures in which higher scores indicate greater symptom severity. The ICIQ-UI SF has a possible range from 0 to 21, the ICIQ-FLUTS filling score ranges from 0 to 16, the ICIQ-FLUTS voiding score ranges from 0 to 12, the ICIQ-FLUTS incontinence score ranges from 0 to 20, the ICIQ-LUTSqol ranges from 19 to 76 and the ICIQ-LUTSqol bother scale ranges from 0 to 10.		

TABLE 3 Pelvic floor measures at baseline

Variable	Treatment group	
	Biofeedback PFMT	Basic PFMT
Prolapse symptoms: POP-SS, <i>n</i> , mean (SD)	274, 6.4 (5.7)	286, 6.7 (5.6)
Oxford Scale: ^a slow contraction strength, <i>n/N</i> (%)		
0	0/300 (0.0)	0/300 (0.0)
1	34/300 (11.3)	31/300 (10.3)
2	115/300 (38.3)	111/300 (37.0)
3	128/300 (42.7)	134/300 (44.7)
4	22/300 (7.3)	24/300 (8.0)
5	1/300 (0.3)	0/300 (0.0)
Oxford Scale: ^a fast contraction strength, <i>n/N</i> (%)		
0	0/238 (0.0)	0/220 (0.0)
1	11/238 (4.6)	14/220 (6.4)
2	75/238 (31.5)	73/220 (33.2)
3	108/238 (45.4)	102/220 (46.4)
4	38/238 (16.0)	28/220 (12.7)
5	6/238 (2.5)	3/220 (1.4)
Contraction endurance [length of hold (seconds)], <i>n</i> , mean (SD)	264, 6.48 (3.00)	250, 6.35 (3.13)
Number of repetitions slow, <i>n</i> , mean (SD)	263, 6.03 (2.44)	249, 5.77 (2.41)
Number of repetitions fast, <i>n</i> , mean (SD)	248, 8.24 (2.50)	239, 7.81 (2.64)
Self-efficacy scale for PFMT, <i>n</i> , mean (SD) ^b	280, 62.7 (9.7)	295, 62.2 (8.8)
Belief in PFMT execution	280, 38.9 (7.1)	295, 38.5 (6.3)
Belief in performing PFMT as scheduled	282, 23.8 (3.8)	294, 23.6 (3.8)
Frequency of PFMT in previous month, <i>n/N</i> (%)		
None	116/287 (40.4)	119/295 (40.3)
Few times a month	37/287 (12.9)	32/295 (10.8)
Once a week	21/287 (7.3)	14/295 (4.7)
Few times a week	52/287 (18.1)	69/295 (23.4)
Once a day	32/287 (11.1)	32/295 (10.8)
Few times a day	29/287 (10.1)	29/295 (9.8)

a The Oxford scale ranges from 0 to 5; higher scores indicate greater strength.

b The self-efficacy scale has a possible range of 17–85 (with the subscales ranging from 11 to 55 and 6 to 30, respectively). Higher scores indicate greater self-efficacy.

in the biofeedback PFMT group. There were some participants for whom compliance status could not be determined due to missing checklist items in the TAFs.

Four participants randomised to the biofeedback PFMT group were given basic PFMT from the outset (in error). Similarly, one participant was randomised to the basic PFMT group but received biofeedback PFMT.

TABLE 4 Appointment attendance and therapist type

Variable	Treatment group	
	Biofeedback PFMT	Basic PFMT
Number of appointments attended, <i>n/N</i> (%)		
0	16/295 (5.4)	33/298 (11.1)
1	20/295 (6.8)	18/298 (6.0)
2	24/295 (8.1)	22/298 (7.4)
3	37/295 (12.5)	33/298 (11.1)
4	33/295 (11.2)	22/298 (7.4)
5	56/295 (19.0)	64/298 (21.5)
6	109/295 (36.9)	106/298 (35.6)
Total number of appointments, <i>n</i> , mean (SD)	295, 4.2 (1.9)	298, 4.0 (2.1)
Type of therapist, <i>n/N</i> (%)		
Physiotherapist	256/295 (86.8)	247/298 (82.9)
Nurse	17/295 (5.8)	11/298 (3.7)
Other and mixture	6/295 (2.0)	7/298 (2.3)
No therapist	16/295 (5.4)	33/298 (11.1)

The exercise programmes are summarised in *Table 5*. The mean recommended length of hold increased from 7.0 seconds in the biofeedback PFMT group and 6.8 seconds in the basic PFMT group to 10.4 seconds and 10.6 seconds, respectively, at the final appointment. The mean number of repetitions increased between the first and last appointment from 7.6 to 8.9 (biofeedback PFMT group) and 6.7 to 9.0 (basic PFMT group). Similarly, the mean number of fast contractions increased from 9.9 to 13.0 (biofeedback PFMT group) and 8.9 to 12.7 (basic PFMT group). The recommended levels for endurance and repetitions at the first appointment appear to be similar to the baseline levels summarised in *Table 3*. *Table 5* also summarises the recommended daily frequency for carrying out the whole set of exercises, which ranged between three and four times a day in both groups.

Primary outcome measure

The ICIQ-UI SF score at 24 months was the primary outcome measure in the trial (lower scores indicate lower symptom severity). In the biofeedback PFMT group, the mean score was 8.2 (SD 5.1) compared with 8.5 (SD 4.9) in the basic PFMT group, with a mean difference (adjusted for baseline score and minimisation covariates) of -0.09 (95% CI -0.92 to 0.75). There was therefore no evidence of a difference between the groups in terms of UI symptoms. Similar results of no difference between groups were found at the 6- and 12-month time points (*Table 6*).

Sensitivity analyses

Analyses of the primary outcome to examine data under differing assumptions relating to non-compliance and missing data all showed very similar results to the primary ITT analysis (*Figure 3*).

The first sensitivity analysis, which tested non-compliance in a CACE analysis (CACE1), treated participants with an indeterminable compliance status as being compliant and the second analysis (CACE2) assumed these participants to be non-compliant. The results of these analyses estimated mean differences of -0.11 (95% CI -1.05 to 0.82) and -0.13 (95% CI -1.20 to 0.95), respectively (see *Figure 3*).

TABLE 5 Home exercise programme recommended

Variable	Appointment number	Treatment group					
		Biofeedback PFMT			Basic PFMT		
		<i>n</i>	Mean	SD	<i>n</i>	Mean	SD
Contraction endurance [length of hold (seconds)]	1	262	7.0	3.9	257	6.8	3.5
	2	238	7.7	4.0	230	7.8	4.9
	3	186	8.3	3.9	187	9.1	7.8
	4	181	9.5	4.5	172	9.6	5.6
	5	150	10.1	5.1	158	10.9	9.9
	6	176	10.4	6.3	172	10.6	6.8
Number of slow contractions (repetitions)	1	261	7.6	6.6	256	6.7	2.6
	2	237	7.9	2.6	228	7.5	2.9
	3	185	8.6	3.0	186	8.0	2.7
	4	180	9.0	3.4	172	8.4	2.6
	5	149	8.9	2.6	159	8.9	2.6
	6	177	8.9	2.9	173	9.0	2.5
Number of fast contractions (repetitions)	1	251	9.9	5.8	251	8.9	4.2
	2	226	10.7	6.2	220	10.3	5.0
	3	173	11.5	6.5	181	11.9	6.3
	4	176	12.5	9.3	164	12.6	7.0
	5	146	12.6	7.4	154	13.2	10.6
	6	174	13.0	10.1	167	12.7	6.6
Number of times per day	1	257	3.2	1.3	255	3.7	1.4
	2	232	3.2	1.3	227	3.7	1.0
	3	183	3.2	1.2	185	3.7	1.2
	4	177	3.1	1.1	171	3.6	1.0
	5	149	3.2	1.1	156	3.8	2.4
	6	174	3.2	2.1	172	3.2	1.3

TABLE 6 Summary of ICIQ-UI SF responses and analysis of differences between groups (models correct for minimisation variables and baseline ICIQ-UI SF)

Time point	Treatment group						Mean difference	95% CI
	Biofeedback PFMT			Basic PFMT				
	<i>n</i>	Mean	SD	<i>n</i>	Mean	SD		
Baseline	291	12.5	4.1	294	12.3	3.7		
6 months	221	9.0	5.0	221	8.8	4.5	0.39	−0.33 to 1.12
12 months	249	9.1	4.9	252	8.7	5.0	0.57	−0.17 to 1.31
24 months	225	8.2	5.1	235	8.5	4.9	−0.09	−0.92 to 0.75

Reproduced with permission from Hagen *et al.*³⁸ This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>. The table includes minor additions and formatting changes to the original table.

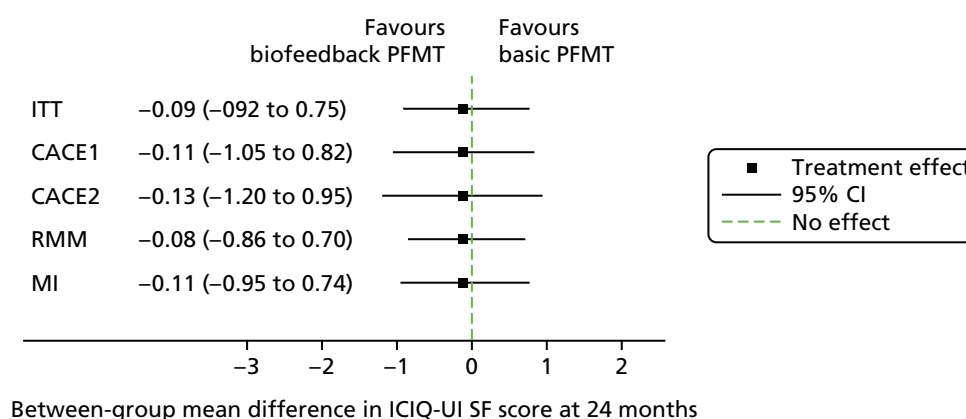


FIGURE 3 Sensitivity analyses. MI, multiple imputation model; RMM, repeated-measures model. Reproduced with permission from Hagen *et al.*³⁸ This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>. The figure includes minor additions and formatting changes to the original figure.

The multiple imputation and the repeated-measures models, which both treat missing data as missing at random, estimated mean differences of -0.11 (95% CI -0.95 to 0.74) and -0.08 (95% CI -0.86 to 0.70), respectively (see *Figure 3*). The repeated-measures model also provided estimates at the intermediate time points, which showed very similar results to the primary analysis: 0.28 (95% CI -0.51 to 1.07) at 6 months and 0.57 (95% CI -0.19 to 1.33) at 12 months. None of the sensitivity analyses performed under MNAR assumptions yielded results that contradicted the primary ITT analysis (see *Appendix 5* and *Figure 12*).

An additional post hoc sensitivity analysis, in which therapist type was removed from the ITT analysis as a covariate in the model (to address the potential bias of including an independent variable that was measured post randomisation), gave a very similar result (mean difference -0.18 , 95% CI -1.02 to 0.65).

Subgroup analyses

The prespecified subgroup analyses of the primary outcome showed no significant treatment by subgroup interactions (*Figure 4*). Full results of the subgroup analysis are included in *Appendix 6*.

Urinary outcomes

The results of the analysis of the lower urinary tract symptoms data are summarised in *Table 7*. The outcome measures used are all International Consultation on Incontinence Questionnaire measures,^{30,31} in which higher scores indicate greater symptom severity or impact on quality of life. There were no significant differences between the groups in any of the outcomes.

Measures of cure and improvement are summarised in *Table 8*. We defined 'cure' as being asymptomatic, identified as a null response to the first two questions of the ICIQ-UI SF (amount and frequency of leakage). Using this definition, 7.9% of participants in the biofeedback PFMT group and 8.4% of participants in the basic PFMT group were asymptomatic at 24 months [odds ratio (OR) 0.90 , 95% CI 0.46 to 1.78]. Three participants in the basic PFMT group were asymptomatic according to this definition at baseline, but they all reported non-zero scores on the ICIQ-UI SF scale for interference of UI with everyday life.

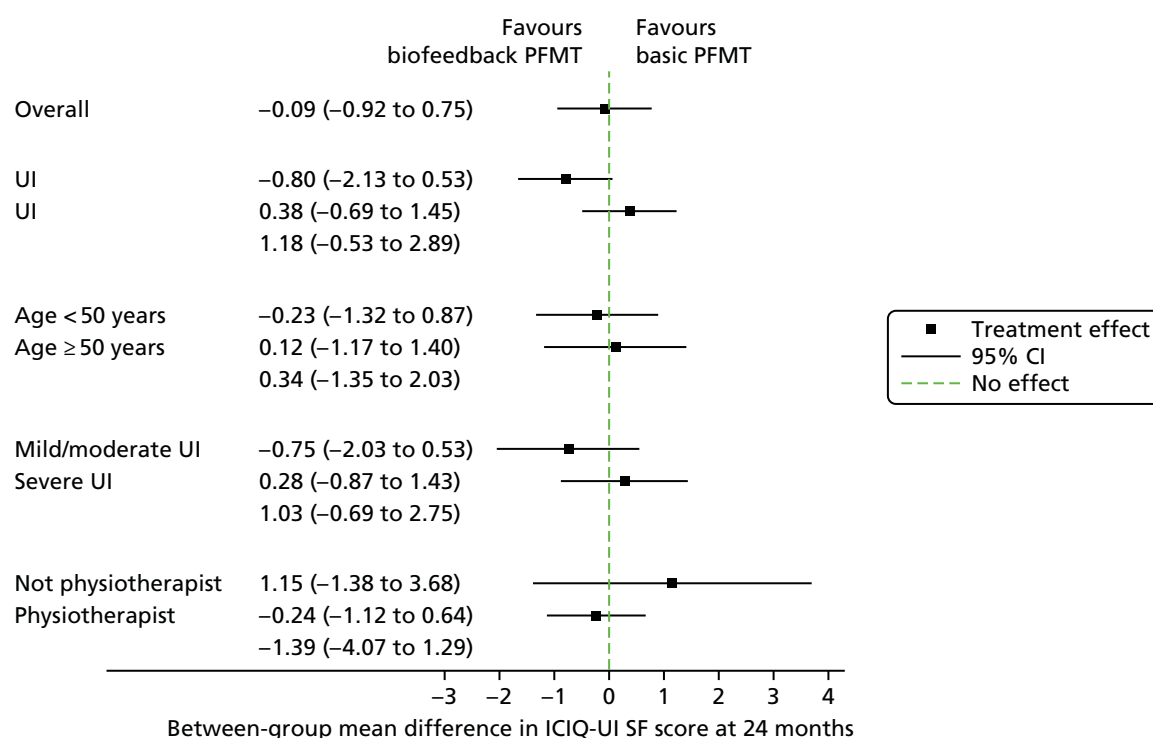


FIGURE 4 Subgroup analyses. Reproduced with permission from Hagen *et al.*³⁸ This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>. The figure includes minor additions and formatting changes to the original figure.

TABLE 7 Lower urinary tract symptoms (ICIQ-FLUTS and ICIQ-LUTSqol)

Variable	Time point	Treatment group						Mean difference at 24 months	95% CI
		Biofeedback PFMT			Basic PFMT				
		n	Mean	SD	n	Mean	SD		
ICIQ-FLUTS filling score	Baseline	289	5.0	2.8	297	4.8	2.6	−0.19	−0.61 to 0.24
	6 months	183	3.7	2.7	176	3.4	2.3		
	12 months	187	3.8	2.7	186	3.6	2.4		
	24 months	167	3.4	2.6	168	3.5	2.3		
ICIQ-FLUTS voiding score	Baseline	292	2.0	2.0	294	2.0	2.1	0.04	−0.30 to 0.38
	6 months	182	1.6	1.8	179	1.4	1.8		
	12 months	188	1.5	1.9	186	1.5	1.8		
	24 months	165	1.6	1.8	169	1.6	1.8		
ICIQ-FLUTS incontinence score	Baseline	290	9.8	3.6	294	9.3	3.4	0.20	−0.58 to 0.98
	6 months	182	7.1	4.0	178	6.6	3.8		
	12 months	188	7.1	3.9	182	6.6	4.1		
	24 months	164	7.0	4.3	169	6.5	4.0		
continued									

continued

TABLE 7 Lower urinary tract symptoms (ICIQ-FLUTS and ICIQ-LUTSqol) (*continued*)

Variable	Time point	Treatment group						Mean difference at 24 months	95% CI
		Biofeedback PFMT			Basic PFMT				
		n	Mean	SD	n	Mean	SD		
ICIQ-LUTSqol	Baseline	292	43.5	12.3	297	42.3	12.1		
	6 months	183	36.2	13.2	176	35.7	11.9		
	12 months	189	35.7	13.3	184	34.7	12.1		
	24 months	164	34.3	12.4	169	34.3	12.5	−0.81	−3.03 to 1.41
ICIQ-LUTSqol bother scale	Baseline	288	7.4	2.6	288	7.6	2.5		
	6 months	183	4.3	3.1	177	4.3	2.8		
	12 months	189	4.0	3.1	184	3.9	3.0		
	24 months	163	3.8	3.1	169	3.7	2.9	0.26	−0.33 to 0.85

Reproduced with permission from Hagen *et al.*³⁸ This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>. The table includes minor additions and formatting changes to the original table.

TABLE 8 Cure and improvement in urinary symptoms

		Treatment group						OR	95% CI
		Biofeedback PFMT			Basic PFMT				
Variable	Time point	<i>n</i>	<i>N</i>	%	<i>n</i>	<i>N</i>	%		
Asymptomatic (cure)	Baseline	0	292	0.0	3	297	1.0		
	6 months	12	221	5.4	13	223	5.8		
	12 months	16	250	6.4	22	253	8.7		
	24 months	18	229	7.9	20	238	8.4	0.90	0.46 to 1.78
Improvement in UI (≥ 3-point reduction in ICIQ-UI SF)	6 months	129	221	58.4	133	221	60.2		
	12 months	148	249	59.4	163	252	64.7		
	24 months	135	225	60.0	147	235	62.6	0.89	0.61 to 1.32
Very much better or much better (PGI-I)	6 months	96	219	43.8	85	221	38.5		
	12 months	101	249	40.6	92	250	36.8		
	24 months	93	227	41.0	90	236	38.1	1.12	0.76 to 1.63

OR, odds ratio.

Reproduced with permission from Hagen *et al.*³⁸ This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>. The table includes minor additions and formatting changes to the original table.

In terms of improvement, 60.0% of participants in the biofeedback PFMT group and 62.6% of participants in the basic PFMT group (OR 0.89, 95% CI 0.61 to 1.32) had a reduction in their ICIQ-UI SF score between baseline and 24 months by a margin of ≥ 3 points (equivalent to the estimated minimum clinically important difference of ≥ 2.5 points).²⁸ Comparing the PGI-I in UI (ordinal scale)²⁹ also shows no difference between groups at 24 months (OR 1.14, 95% CI 0.75 to 1.72). Full descriptive summaries of the PGI-I data are reported (see *Appendix 7*). As a post hoc analysis, data from a dichotomous version of this scale were compared between groups (see *Table 8*). At 24 months, 41.0% of participants in the biofeedback PFMT group and 38.1% of participants in the basic PFMT group reported being much better or very much better (OR 1.12, 95% CI 0.76 to 1.63), which resulted in a very similar effect estimate to the analysis of the full ordinal PGI-I data (OR 1.14, 95% CI 0.75 to 1.72). When tested, the proportional odds assumption held in the ordinal analysis.

Further treatment for urinary incontinence

Participants who accessed other treatments for UI (including surgery) were able to remain in the trial. *Table 9* summarises the rates of surgical and non-surgical treatment for UI reported by participants. In the first 6 months post randomisation, 1.2% of women in the biofeedback PFMT group and 1.8% of women in the basic PFMT group had continence surgery (OR 0.56, 95% CI 0.09 to 3.53). This increased to 3.9% and 5.2%, respectively, in the second 6 months (OR 0.63, 95% CI 0.23 to 1.69) and to 5.2% and 7.4%, respectively, during the second year of follow-up (OR 0.62, 95% CI 0.24 to 1.65). The cumulative rate for UI surgery over the whole 2 years was 12.3% in the biofeedback PFMT group and 9.3% in the basic PFMT group (OR 1.25, 95% CI 0.35 to 4.46).

Rates of non-surgical treatments did not increase over the course of follow-up. There were high cumulative rates of uptake over the 2-year follow-up period in both groups (81.7%, biofeedback PFMT; 79.7%, basic PFMT, OR 1.35, 95% CI 0.54 to 3.41), although this measure includes a broad range of health-care use for the treatment of UI, some resources of which were accessed frequently. For example, appointments with either a continence nurse in secondary care or a physiotherapist were attended by 60.6% in the biofeedback PFMT group and 53.8% in the basic PFMT group (*Table 10*).

TABLE 9 Further treatment for UI

		Treatment group						OR	95% CI
		Biofeedback PFMT			Basic PFMT				
Treatment	Time point	<i>n</i>	<i>N</i>	%	<i>n</i>	<i>N</i>	%		
Surgery	0–6 months	2	172	1.2	3	164	1.8	0.56	0.09 to 3.53
	6–12 months	8	204	3.9	11	210	5.2	0.63	0.23 to 1.69
	12–24 months	8	154	5.2	12	162	7.4	0.62	0.24 to 1.65
	0–24 months	10	81	12.3	8	86	9.3	1.25	0.35 to 4.46
Non-surgical treatment	0–6 months	96	146	65.8	107	149	71.8	0.77	0.46 to 1.28
	6–12 months	70	164	42.7	74	159	46.5	0.90	0.56 to 1.42
	12–24 months	40	105	38.1	42	119	35.3	0.65	0.65 to 2.03
	0–24 months	49	60	81.7	55	69	79.7	1.35	0.54 to 3.41

Reproduced with permission from Hagen *et al.*³⁸ This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>. The table includes minor additions and formatting changes to the original table.

TABLE 10 Hospital admissions and non-surgical treatment for UI

Treatment	Time point	Treatment group	
		Biofeedback PFMT	Basic PFMT
Admitted to hospital, <i>n/N</i> (%)	0–6 months	3/180 (1.7)	2/179 (1.1)
	6–12 months	4/180 (2.2)	4/175 (2.3)
	12–24 months	3/111 (2.7)	1/130 (0.8)
	0–24 months	6/84 (7.1)	3/94 (3.2)
Number of nights in hospital, <i>n</i> , mean (SD)	0–6 months	180, 0.02 (0.13)	179, 0.01 (0.11)
	6–12 months	180, 0.06 (0.54)	175, 0.03 (0.26)
	12–24 months	111, 0.04 (0.23)	130, 0.01 (0.09)
	0–24 months	84, 0.16 (0.81)	94, 0.05 (0.34)
Hospital doctor appointment, <i>n/N</i> (%)	0–6 months	43/156 (27.6)	46/163 (28.2)
	6–12 months	27/185 (14.6)	35/181 (19.3)
	12–24 months	18/124 (14.5)	20/131 (15.3)
	0–24 months	35/83 (42.2)	34/89 (38.2)
GP appointment, <i>n/N</i> (%)	0–6 months	46/183 (25.1)	65/183 (35.5)
	6–12 months	32/187 (17.1)	29/186 (15.6)
	12–24 months	18/125 (14.4)	21/132 (15.9)
	0–24 months	37/99 (37.4)	48/100 (48.0)
GP nurse appointment, <i>n/N</i> (%)	0–6 months	13/180 (7.2)	6/180 (3.3)
	6–12 months	16/178 (9.0)	11/183 (6.0)
	12–24 months	8/127 (6.3)	6/128 (4.7)
	0–24 months	14/96 (14.6)	10/94 (10.6)
Hospital nurse/physiotherapist appointment, <i>n/N</i> (%)	0–6 months	62/177 (35.0)	57/179 (31.8)
	6–12 months	51/181 (28.2)	49/181 (27.1)
	12–24 months	14/124 (11.3)	21/129 (16.3)
	0–24 months	57/94 (60.6)	50/93 (53.8)
Medication, <i>n/N</i> (%)	0–6 months	32/209 (15.3)	34/204 (16.7)
	6–12 months	32/234 (13.7)	38/232 (16.4)
	12–24 months	30/168 (17.9)	28/174 (16.1)
	0–24 months	35/125 (28.0)	36/128 (28.1)
Other treatment/advice, <i>n/N</i> (%)	0–6 months	25/208 (12.0)	29/201 (14.4)
	6–12 months	22/229 (9.6)	27/232 (11.6)
	12–24 months	12/208 (5.8)	15/221 (6.8)
	0–24 months	35/157 (22.3)	35/162 (21.6)

Pelvic floor outcomes

When self-efficacy for PFMT³⁴ was compared, there was a small and statistically significant difference in favour of the biofeedback PFMT group in the overall score (mean difference 2.36, 95% CI 0.04 to 4.68) (*Table 11*). This difference appeared to be attributed more to the subscale for belief in PFMT and its execution, in which there was also a small and statistically significant difference (mean difference 1.54, 95% CI 0.03 to 3.06). There was no significant difference in the subscale for belief in performing PFMT as scheduled and despite barriers (mean difference 0.71, 95% CI –0.31 to 1.72).

TABLE 11 Self-efficacy for PFMT

Subgroup	Time point	Treatment group						Mean difference	95% CI
		Biofeedback PFMT			Basic PFMT				
		n	Mean	SD	n	Mean	SD		
Self-efficacy scale for PFMT	Baseline	280	62.7	9.7	295	62.2	8.8		
	6 months	172	65.6	10.2	165	65.7	10.1		
	12 months	178	63.7	11.0	180	64.1	10.2		
	24 months	154	63.1	11.6	161	60.9	12.0	2.36	0.04 to 4.68
Belief in PFMT execution subscale	Baseline	280	38.9	7.1	295	38.5	6.3		
	6 months	173	42.7	7.0	166	42.3	6.8		
	12 months	179	41.7	7.3	180	41.6	6.6		
	24 months	156	41.3	7.8	162	39.8	7.9	1.54	0.03 to 3.06
Belief in performing PFMT as scheduled subscale	Baseline	282	23.8	3.8	294	23.6	3.8		
	6 months	178	23.0	4.3	172	23.5	4.3		
	12 months	183	21.9	4.9	179	22.6	4.8		
	24 months	156	21.6	4.9	164	21.0	5.0	0.71	−0.31 to 1.72

The results from the 6-month blinded pelvic floor muscle assessment are summarised in *Table 12*. At baseline, there was only one participant who achieved the maximum score of 5 on the Oxford Scale for slow contraction strength. At 6 months, there were 13 participants (8.5%) in the biofeedback PFMT group and 10 (6.0%) in the basic PFMT group with a score of 5. There was no difference between the groups when the Oxford Scale at 6 months was compared (OR 1.28, 95% CI 0.86 to 1.89; $p = 0.22$). The proportional odds assumption, however, did not hold in this ordinal analysis.

TABLE 12 Pelvic floor muscle function assessed by therapist via vaginal examination at baseline and 6 months (blinded)

Variable	Time point	Score	Treatment group	
			Biofeedback PFMT	Basic PFMT
Oxford scale: slow contraction strength, n/N (%)	Baseline	0	0/300 (0.0)	0/300 (0.0)
		1	34/300 (11.3)	31/300 (10.3)
		2	115/300 (38.3)	111/300 (37.0)
		3	128/300 (42.7)	134/300 (44.7)
		4	22/300 (7.3)	24/300 (8.0)
		5	1/300 (0.3)	0/300 (0.0)
	6 months	0	0/153 (0.0)	0/166 (0.0)
		1	4/153 (2.6)	3/166 (1.8)
		2	25/153 (16.3)	23/166 (13.9)
		3	57/153 (37.3)	74/166 (44.6)
		4	54/153 (35.3)	56/166 (33.7)
		5	13/153 (8.5)	10/166 (6.0)

continued

TABLE 12 Pelvic floor muscle function assessed by therapist via vaginal examination at baseline and 6 months (blinded) (*continued*)

Variable	Time point	Score	Treatment group	
			Biofeedback PFMT	Basic PFMT
Oxford scale: fast contraction strength, <i>n/N</i> (%)	Baseline	0	0/238 (0.0)	0/220 (0.0)
		1	11/238 (4.6)	14/220 (6.4)
		2	75/238 (31.5)	73/220 (33.2)
		3	108/238 (45.4)	102/220 (46.4)
		4	38/238 (16.0)	28/220 (12.7)
		5	6/238 (2.5)	3/220 (1.4)
	6 months	0	1/133 (0.8)	0/152 (0.0)
		1	2/133 (1.5)	2/152 (1.3)
		2	30/133 (22.6)	22/152 (14.5)
		3	49/133 (36.8)	66/152 (43.4)
		4	39/133 (29.3)	49/152 (32.2)
		5	12/133 (9.0)	13/152 (8.6)
Contraction endurance [length of hold (seconds)], <i>n</i> , mean (SD)	Baseline		264, 6.48 (3.00)	250, 6.35 (3.13)
	6 months		152, 8.72 (2.26)	166, 8.54 (2.48)
Number of slow contractions (repetitions), <i>n</i> , mean (SD)	Baseline		263, 6.03 (2.44)	249, 5.77 (2.41)
	6 months		151, 7.42 (2.62)	165, 7.55 (2.59)
Number of fast contractions (repetitions), <i>n</i> , mean (SD)	Baseline		248, 8.24 (2.50)	239, 7.81 (2.64)
	6 months		149, 8.94 (2.14)	164, 9.50 (1.50)

For endurance and repetitions, the results at 6 months appear to be similar between groups and appear to be slightly higher than baseline in both groups, although it should be noted that nearly half of the participants did not attend the 6-month pelvic floor muscle appointment. Furthermore, at 6 months, the observed results for endurance and repetitions appear to be slightly lower than the prescribed programme at the end of the intervention period (see *Tables 5 and 12*).

Bowel and prolapse symptoms are summarised at each time point (see *Appendix 8*). No validated measure for bowel symptoms was available, but the results are summarised for each question asked individually (see *Table 35*). There was no evidence of any effect on prolapse symptoms with no difference between groups in the POP-SS (mean difference -0.60 , 95% CI -1.51 to 0.30) (see *Table 36*).

Adherence

Our prespecified definitions for adherence are described in the statistical analysis plan [see the project web page: www.journalslibrary.nihr.ac.uk/programmes/hta/117103/#/ (accessed 29 July 2019)]. *Table 13* summarises the levels of adherence to separate aspects of the intervention; there were no differences between groups. Adherence to introductory teaching of PFMT and introduction to biofeedback (as appropriate) by the therapist was 88% in each group (OR 0.69, 95% CI 0.33 to 1.42). Adherence to practising PFMT and biofeedback use during appointments was just under 80% in each group (OR 0.89, 95% CI 0.63 to 1.25). Adherence by participants to the recommended programme at home between appointments was also around 80% in each group (OR 0.71, 95% CI 0.43 to 1.16). *Table 13* also summarises the frequency of appointments

TABLE 13 Therapists' adherence to intervention delivery protocol at appointments and women's adherence to prescribed intervention at home

Variable	Treatment group						OR	95% CI
	Biofeedback PFMT				Basic PFMT			
	<i>n</i>	<i>N</i>	%	<i>n</i>	<i>N</i>	%		
Adherence during clinic appointment								
Introductory teaching (as appropriate to group)	254	288	88.2	259	293	88.4	0.69	0.33 to 1.42
Any adherence in clinic	231	290	79.7	231	292	79.1	0.89	0.63 to 1.25
Frequency of adherence to intervention protocol at clinic appointments (subsequent to initial appointment)								
One appointment or more	218	277	78.7	200	261	76.6		
Two appointments or more	175	277	63.2	167	261	64.0		
Three appointments or more	130	277	46.9	121	261	46.4		
Four appointments or more	86	277	31.0	89	261	34.1		
All five appointments	39	277	14.1	40	261	15.3		
Adherence at home								
Any adherence at home	220	281	78.3	241	297	81.1	0.71	0.43 to 1.16
Frequency of adherence to intervention protocol at home (number of appointments with adherence since previous appointment)								
Once or more	196	257	76.3	177	233	76.0		
Twice or more	155	257	60.3	156	233	67.0		
Three times or more	115	257	44.7	125	233	53.6		
Four times or more	75	257	29.2	105	233	45.1		
All five times	34	257	13.2	52	233	22.3		

with intervention adherence (subsequent to the first appointment) and the frequency of adherence at home (between appointments). There were some participants for whom adherence status could be determined, but the frequency of adherence could not be determined as a result of partially missing checklist data in the TAFs. There was also one participant in the biofeedback PFMT group and four in the basic PFMT group who, at 24 months, reported using a biofeedback device at home after the OPAL trial intervention had completed.

The frequency of PFMT being undertaken, as reported by participants across the different post-intervention time points, is summarised in *Table 14*. At 24 months, the proportion of participants exercising at least once a week was 52.0% in the biofeedback PFMT group and 46.3% in the basic PFMT group.

Adverse events

There were no suspected unexpected serious adverse reactions reported by participants in the trial. There were eight SAEs (six in the biofeedback PFMT group and two in the basic PFMT group), all of which were unrelated to the intervention received (e.g. two participants had atrial fibrillation).

In addition to the SAEs, there were 48 participants for whom at least one AE was reported (34 in the biofeedback PFMT group and 14 in the basic PFMT group), but none of these complications was classified as a SAE. Of these participants, 23 (21, biofeedback PFMT; 2, basic PFMT) had a complication related or possibly related to one of the interventions. Only one of these complications was definitely related to the intervention, when a participant randomised to the biofeedback PFMT group was found to have a nickel allergy and did not continue with the intervention.

TABLE 14 Participant-reported frequency of PFMEs

Time point	Time frame	Frequency	Treatment group	
			Biofeedback PFMT	Basic PFMT
At 6 months	Yesterday, <i>n/N</i> (%)	None	47/167 (28.1)	46/166 (27.7)
		A little	47/167 (28.1)	42/166 (25.3)
		Now and then	32/167 (19.2)	17/166 (10.2)
		Regularly	41/167 (24.6)	61/166 (36.7)
	In previous 7 days, <i>n</i> , mean (SD)	Number of days	168, 4.5 (2.0)	167, 5 (2.0)
	Over previous month, <i>n/N</i> (%)	None	19/183 (10.4)	10/172 (5.8)
		A few times a month	11/183 (6.0)	10/172 (5.8)
		Once a week	12/183 (6.6)	8/172 (4.7)
		A few times a week	36/183 (19.7)	24/172 (14.0)
		Once a day	31/183 (16.9)	39/172 (22.7)
		A few times a day	74/183 (40.4)	81/172 (47.1)
At 12 months	Yesterday, <i>n/N</i> (%)	None	59/147 (40.1)	59/156 (37.8)
		A little	47/147 (32.0)	51/156 (32.7)
		Now and then	19/147 (12.9)	21/156 (13.5)
		Regularly	22/147 (15.0)	25/156 (16.0)
	In previous 7 days, <i>n</i> , mean (SD)	Number of days	145, 4 (2.2)	153, 4.1 (2.3)
	Over previous month, <i>n/N</i> (%)	None	52/197 (26.4)	40/195 (20.5)
		A few times a month	20/197 (10.2)	24/195 (12.3)
		Once a week	12/197 (6.1)	11/195 (5.6)
		A few times a week	51/197 (25.9)	41/195 (21.0)
		Once a day	27/197 (13.7)	43/195 (22.1)
		A few times a day	35/197 (17.8)	36/195 (18.5)
At 24 months	Yesterday, <i>n/N</i> (%)	None	64/132 (48.5)	70/128 (54.7)
		A little	39/132 (29.5)	31/128 (24.2)
		Now and then	12/132 (9.1)	10/128 (7.8)
		Regularly	17/132 (12.9)	17/128 (13.3)
	In previous 7 days, <i>n</i> , mean (SD)	Number of days	133, 3.5 (2.3)	129, 3.5 (2.4)
	Over previous month, <i>n/N</i> (%)	None	47/173 (27.2)	63/188 (33.5)
		A few times a month	36/173 (20.8)	38/188 (20.2)
		Once a week	5/173 (2.9)	7/188 (3.7)
		A few times a week	45/173 (26.0)	31/188 (16.5)
		Once a day	17/173 (9.8)	27/188 (14.4)
		A few times a day	23/173 (13.3)	22/188 (11.7)

Chapter 4 Health economic evaluation

Methods

The principal research question being addressed in the economic analysis concerns the cost-effectiveness of a policy of biofeedback PFMT compared with basic PFMT. The main economic evaluation was based on data collected alongside the RCT. There was a plan to undertake an additional modelling analysis that considered a longer time horizon, to provide additional information for policy-makers, if relevant. However, the evidence indicated that there were no statistical or clinically meaningful differences in either the clinical or the economic outcomes to warrant further extrapolation of the data. Therefore, no modelling was conducted and this chapter focusses on the within-trial analysis. The trial population were women presenting with SUI or MUI. The base-case trial analysis assessed the costs and cost-effectiveness of the interventions from the perspective of the NHS. A further analysis included a societal perspective that considered the cost to the participants and their families. The costs were in Great British pounds (GBP) and the cost year was 2017. All analyses were prespecified in the health economic analysis plan [see the project web page: www.journalslibrary.nihr.ac.uk/programmes/hta/117103/#/ (accessed 29 July 2019)] and were reported following the Consolidated Health Economic Evaluation Reporting Standards (CHEERS) recommendations [URL: www.equator-network.org/reporting-guidelines/cheers/ (accessed 29 July 2019)]. The methods are summarised below.

Data collection

Resource utilisation was measured over four time points (baseline and 6, 12 and 24 months) using TAFs and participant-completed questionnaires, which included health-care utilisation questions. The second year costs and benefits were discounted using the National Institute for Health and Care Excellence (NICE) recommended 3.5%.³⁹ Intervention resource use included the number of appointments with the therapists as captured on the TAF. Although the first appointment was expected to last for 1 hour and each follow-up appointment to last for 30 minutes, data on the length of each appointment were also recorded on the TAF. The biofeedback PFMT group had the additional resource of biofeedback in the clinic and a portable biofeedback unit for home use. Information on the biofeedback units, such as type of units, type and number of electrodes, internal probes, laptops, software and licences for laptops, printers and cables, was sourced from the trial office.

Further resource use relating to UI was recorded retrospectively for every woman in the trial. These data were collected at 6, 12 and 24 months, using participant-completed questionnaires and included the use of primary (GP services) and secondary (outpatient appointments, inpatient stay and surgical interventions for incontinence) health care. Health service costs refer to those incurred directly by the NHS, such as related surgery, additional appointments and procedures. As physiotherapist appointments were part of the trial interventions, it was probable that participant-reported appointments in the 6-month questionnaire included some or all of the trial-related appointments and would introduce the chance of double-counting number and cost. Therefore, the numbers of trial appointments and participant-reported incontinence-related appointments were compared to establish a clear pattern in the latter, in an attempt to exclude any already captured in the intervention costing. It was assumed that a one-appointment difference in participant-reported appointments was most probably due to patient recall bias. It was assumed that these participants had no additional appointments, other than the trial appointments, and incurred no additional cost. Those women who reported more than one appointment in addition to their trial appointments incurred the cost of the difference in participant-reported and trial appointments. Resource use incurred at personal cost to the participants (such as purchase of pads, medication) was collected using the participant-reported 24-month questionnaire.

Intervention costs

The cost of the biofeedback unit was calculated on a per-participant basis and was estimated based on the following assumptions: the cost of biofeedback units and equipment (laptops, printers, cables) obtained from the trial office was spread out over 5 years, reflecting the expected lifespan of all equipment (Amanda Tombs, Verity Medical Ltd, 2018, personal communication). It was assumed that, on average, three women would use the unit every year to allow for the 16-week intervention period and time for units to be returned to the centres. The cost of consumables (two packets of electrodes and a probe) was also included. As the biofeedback units and consumables were purchased across several years, the Hospital and Community Health Service (HCHS) index⁴⁰ was used to bring all costs to the same year. The cost of the biofeedback unit per user was then annuitised, using an annual rate of 3.5% over its lifespan of 5 years. This produced a per-participant unit cost of biofeedback of £37.40 for the 16-week intervention period (see *Appendix 9* for details on the costing).

Medication costs

The incontinence care pathway suggests that, before women progress to the severity level, when surgery is needed, treatment with medications for UI and overactive bladder is recommended.¹⁰ Incontinence medications have a very low unit cost; Mirabegron (Betmiga; Astellas Pharma Inc., Tokyo, Japan) and Solifenacin (Vesicare; Astellas Pharma Inc.) are the most frequently reported and have the highest unit cost (see *Appendix 10*). As participant-reported medication data were available only at distinct time points (from 6-, 12- and 24-month follow-up questionnaires), an assumption was made that participants started their treatment halfway between the follow-up points.

Unit costs

Unit costs and prices were obtained using published estimates from the *British National Formulary* for medications,⁴¹ NHS Reference Costs 2016/17⁴² for secondary care resource use and the Personal Social Services Research Unit's *Unit Costs of Health and Social Care 2017*⁴⁰ for primary care resource use, as outlined in *Table 15*.

TABLE 15 Average NHS unit costs

Area of resource use	Resource	Unit cost (£)	Source	Notes
Intervention resource use	Portable biofeedback units	115	Trial office	Average cost of biofeedback units purchased in 2013–15 and inflated to 2017 prices using HCHS
	Electrodes	8	Trial office	Average cost of electrodes purchased for the trial in 2013–14 and inflated to 2017 prices using HCHS
	Probe	13	Trial office	Average cost of internal probes purchased for the trial in 2013–15 and inflated to 2017 prices using HCHS
	Average unit cost per unit	37	Estimated cost	Based on information outlined in <i>Intervention costs</i> . See <i>Appendix 10</i> for details
	Therapist appointment (physiotherapist or specialist nurse)	2	Curtis and Burns 2017 ⁴⁰	Per-minute cost. Based on average cost per hour of patient contact of band 6 and band 7 hospital-based scientific and professional staff
Primary care	GP appointment	31 (37)	Curtis and Burns 2017 ⁴⁰	Cost per surgery consultation lasting 9.22 minutes, including qualifications given in brackets, both are including direct-care staff costs
	Nurse appointment (general practice)	9 (11)	Curtis and Burns 2017 ⁴⁰	Surgery consultation based on the 2006/7 UK general practice survey is 15.5 minutes (including qualifications given in brackets)

TABLE 15 Average NHS unit costs (*continued*)

Area of resource use	Resource	Unit cost (£)	Source	Notes
Secondary care (outpatient services)	NHS doctor appointment	130 (103)	<i>Reference Costs 2016/17</i> ⁴²	First non-admitted face-to-face appointment (follow-up appointments in brackets), consultant led (urology)
	NHS nurse appointment/physiotherapist	83	<i>Reference Costs 2016/17</i> ⁴²	Specialist nursing, continence services, adult, face to face
Secondary care (inpatient)	Overnight stay in hospital	399	<i>Reference Costs 2016/17</i> ⁴²	Weighted average for non-elective short-term stay (0/1 nights), for UI without interventions across all severity scores (HRG codes LB16G-K)
		1724	<i>Reference Costs 2016/17</i> ⁴²	Weighted average for non-elective long-term stay (2/3 nights) for UI without interventions across all severity scores (HRG codes LB16G-K)
		278	<i>Reference Costs 2016/17</i> ⁴²	Excess bed-days (≥ 4 nights) for UI without interventions across all severity scores (HRG codes LB16G-K)
Participant resource use	Medications	Various	<i>British National Formulary</i> ⁴¹	Participant reported. See <i>Appendix 10</i> for more detail
	Surgical interventions	1680	<i>Reference Costs 2016/17</i> ⁴²	A weighted average of HRG4 codes LB51A and LB51B CC score 0–2+ for TVT
	Non-surgical interventions (Injections)	963	<i>Reference Costs 2016/17</i> ⁴²	Intermediate endoscopic bladder procedures (HRG code LB14Z)
Private care	Doctor (GP)	70	Bupa 2018 ⁴³	Based on 15-minute appointment
	Nurse	51 (42)	Bupa 2018 ⁴³	Initial consultation (follow-up consultations in brackets), pay-as-you-go appointments (assumed equivalence with physiotherapist)
	Physiotherapist	51 (42)	Bupa 2018 ⁴³	Initial consultation (follow-up consultations in brackets), pay-as-you-go appointments

CC, complication or comorbidity; HRG, Healthcare Resource Group; TVT, tension-free vaginal tape.

Estimation of participant costs

For each area of resource use, estimates of resource utilisation were multiplied by unit costs to derive total costs for each item of resource use and each participant. These data were averaged to provide estimates of the average cost per participant for each item of resource use. These costs were summed to produce a total cost for each participant and an average total cost per participant in each intervention group.

Participant- and companion-incurred costs and indirect costs

Personal costs to the participants (such as costs of travelling to appointments and work/social restrictions) were also investigated. Participant resource utilisation comprised three main elements: (1) self-purchased health care; (2) travel costs for making return appointment(s) to receive NHS health care (such as petrol, public transport and parking); and (3) time costs of travelling and attending NHS health care (such as time involved away from usual activities or work). The self-purchased health-care items included in the investigations were services, medications and pads paid for by participants for the treatment or management of UI symptoms, and time and travel costs related to time spent travelling to, and attending, hospital or primary care providers in relation to UI. Estimation of travel costs included information from participants about the number of appointments with,

for example, their GP or physiotherapist (estimated from the health-care utilisation questions at the various follow-up questionnaire time points) and the unit cost of making a return journey to each type of health-care provider (from the participant time and travel cost questions collected at 24 months). The cost of participant time was estimated in a similar manner. The participant was asked how long they had spent travelling to, and attending, their previous appointment with each type of health-care provider. Information was sought on the activity they would have been undertaking (e.g. paid work, leisure, housework) had they not attended the health-care provider. They were also asked if they were accompanied by a friend or a relative and their time and travel costs were also incorporated into the analysis. These data were presented in their natural units (e.g. hours) and also costed using standard economic conventions, using the Department for Transport⁴⁴ estimates for the value of work and leisure time. These unit time costs were combined with the number of health-care contacts derived from the health-care utilisation questions, to elicit a total time and travel cost from a participant perspective. Details of unit costs applied to the various activities are included in *Table 16*.

Quality-of-life measures

A generic instrument, the EQ-5D-3L, was used to measure the participants' quality of life. The EQ-5D-3L was completed at baseline and at 6, 12 and 24 months post randomisation. The responses to the EQ-5D-3L questionnaires were valued using UK general population tariffs, based on the time trade-off technique to generate a utility score for every participant in the trial.⁴⁷ QALYs were calculated by combining utility with health state duration, using an area beneath the curve approach, assuming linear interpolation of utility between time points. Quality-of-life data were also collected using items from the condition-specific tool ICIQ-LUTSqol for comparison. This tool was included as it was anticipated that the generic measure (EQ-5D-3L) may not be sensitive enough to pick up the continence-related factors that influence quality of life for these participants, such as impact on personal relationship and sleep.⁴⁸ The ICIQ-LUTSqol is a 21-item, condition-specific measure of health for urinary incontinent participants, developed in a UK population similar to that in this trial. This instrument was revised into a five-dimensional health state classification amenable to valuation, using items selected using psychometric evidence. Forty-nine states were valued using standard gamble by a representative sample of patients with UI attending UK hospital outpatient clinics. These data were converted into a utility index using a published algorithm⁶ and used as an alternative utility value for condition-specific QALY calculation.

TABLE 16 Participant time and travel unit cost

Activity	Unit cost (£)	Source and notes
Unit costs applied to participant and companion travel		
Cost per mile travelled by car	0.45	HMRC ⁴⁵
Car parking charges	Various	As reported by participants
Cost of public transport (bus, train, taxi)	Various	As reported by participants
Unit costs applied to participant and companion time		
Paid work	15.28	ONS annual survey of hours and earnings ⁴⁶
Housework	8.58	ONS annual survey of hours and earnings ⁴⁶
Child care	15.28	ONS annual survey of hours and earnings (as paid work) ⁴⁶
Caring for a friend/family member	15.28	ONS annual survey of hours and earnings (as paid work) ⁴⁶
Voluntary work	15.28	ONS annual survey of hours and earnings (as paid work) ⁴⁶
Retired	5.93	Department for Transport ⁴⁴
Leisure	5.93	Department for Transport ⁴⁴
Unemployed	5.93	Department for Transport ⁴⁴
Private doctor appointments	70.00	Bupa 2018 ⁴³
Private appointments with nurse/physiotherapist	51.00	Bupa 2018 ⁴³
HMRC, HM Revenue and Customs; ONS, Office for National Statistics.		

Data analysis

The economic analysis was undertaken on the ITT principle. All components of costs were described with the appropriate descriptive statistics: mean and SD for continuous and count outcomes; numbers and percentages for dichotomous and categorical outcomes (e.g. numbers reporting problems on EQ-5D-3L). All analyses were conducted using Stata® version 14.1 software (StataCorp LP, College Station, TX, USA). Investigations were carried out for skewed cost data (i.e. a small proportion of participants incurring very high costs), using GLMs to test alternative model specifications for appropriate fit to the data. The GLM allowed for heteroscedasticity by selecting and specifying an appropriate distributional family and link function for the data. This family offered alternative specifications to reflect the relationship between the mean and variance of the estimates under consideration.^{49,50} Two diagnostic actions were performed to identify the most appropriate distributional family: (1) a modified Park test and (2) the Akaike information criterion were consulted. The best model fit for analysing cost data appeared to be the gamma distribution with an identity link. The Gaussian distribution with an identity link was chosen for the QALY analysis. The appropriate link to characterise the relation between the linear combination of predictors and the prediction was selected based on the results from the Pearson correlation, Pregibon link and the modified Hosmer–Lemeshow tests (see *Appendix 12*).

Both cost and QALY difference analyses were adjusted for baseline prognostic factors (all of which were minimisation covariates plus the type of therapist and baseline EQ-5D-3L utility score):

- centre number
- age in years
- type of UI
- UI severity
- type of therapist (physiotherapist/nurse)
- baseline EQ-5D-3L utility score.

The first five factors are in line with the clinical effectiveness analyses and the baseline EQ-5D-3L was included for the economic analysis. Standard parametric tests for differences in costs were performed, with the robustness of the parametric tests confirmed using bias-corrected, non-parametric bootstrapping.⁵¹

Missing data

Missing data are a frequent problem in economic evaluations undertaken in a RCT setting. There are several possible methods that can be employed to account for such missing data, including mean or multiple imputation. Multiple imputation analysis was conducted for the base-case analysis, as > 5% of the quality-of-life and cost data needed were missing for the primary analysis. Components of cost data were imputed, based on linear regression models that were adjusted for minimisation variables, baseline utility and intervention allocation group. Missing utility values were imputed using predictive mean matching (the mean of five nearest values). Chained equations were used for the imputations; 20 imputations were considered sufficient to generate stable and reliable estimates for analysis. Imputed data were then analysed using the appropriately specified GLMs described in *Data analysis*.

Incremental cost per quality-adjusted life-year gained

Incremental cost-effectiveness ratios were computed using the imputed data. Differences in costs were based on resource utilisation costs over the 24-month follow-up period. The difference in effectiveness was expressed in terms of QALYs at 24-month follow-up. The economic results were reported using a cost–utility analysis framework, with the ICER representing the incremental cost per QALY gained. The point estimate of the ICER was calculated as:

$$\text{ICER} = \frac{C_i - C_j}{E_i - E_j} = \frac{\Delta C}{\Delta E}, \quad (1)$$

in which C_i and C_j are the mean costs among participants in the biofeedback PFMT and basic PFMT group, respectively. Similarly, E_i and E_j are the mean QALYs in the biofeedback PFMT and basic PFMT groups,

respectively. The ICER was assessed against the NICE-recommended cost-effectiveness threshold of £20,000–30,000 per QALY gained.

Measures of variance for NHS costs and QALYs were derived using non-parametric bootstrapping. From the results of the bootstrapping, cost-effectiveness acceptability curves (CEACs) were created. CEACs are used to display the inherent uncertainty surrounding cost-effectiveness at various threshold values for society's willingness to pay (WTP) for an additional QALY. CEACs present results when the analysis follows a net benefit approach. This approach utilises a straightforward rearrangement of the cost-effectiveness decision rule used when calculating ICERs to create the net monetary benefit (NMB) for each bootstrapped iteration at increasing values of WTP per QALY:

$$\text{NMB} = \lambda \times \Delta E - \Delta C > 0, \quad (2)$$

in which λ represents a decision-maker's WTP for incontinence avoided or a QALY gained. If the above expression holds true for a given iteration and threshold WTP value (λ), then the intervention is considered cost-effective for that iteration. As society's WTP is unknown, the NMB will be calculated for a number of possible λ values, including the usual £20,000–30,000 range often adopted by policy-makers in the NHS.³⁹ A balance sheet approach was used to report the costs and QALYs of women who are or are not incontinent.

Sensitivity and subgroup analyses

Sensitivity analyses were performed to gauge the impact of varying key assumptions and/or parameter values in the base-case analysis. Sensitivity analyses in relation to the assumptions made in the cost of the intervention were performed. The base-case analysis utilised number of minutes recorded by therapists for each appointment, to estimate the average cost per intervention. The first sensitivity analysis performed was based on the recommended time of appointments of 1 hour for the first appointment and 30 minutes for the following five appointments.¹ The base-case analysis in terms of utilities was adjusted for baseline EQ-5D-3L values to account for variability that may be present among the intervention groups. An unadjusted analysis was also performed as a sensitivity analysis to highlight the importance of this base-case assumption. An analysis exploring the impact of changing the discount rate used for second-year costs and QALYs in accordance with NICE best-practice recommendations, varying the discount rate from 0% to 6% per annum, was undertaken.¹ A subgroup analysis similar to that described in the statistical analysis plan was undertaken. This was based on:

- type of incontinence (SUI or MUI)
- age (< 50/≥ 50 years)
- UI severity (ICIQ-UI SF score of < 13 or ≥ 13).

A subgroup analysis on the type of therapist was not performed, as almost all participants were seen by a physiotherapist.

Sensitivity analysis was also performed to assess the impact of the assumption that data were MNAR. Several methods have been proposed in the statistical literature to conduct analyses under MNAR.⁵² The analysis was conducted using pattern mixture models to capture how the distribution of missing values could differ from the conditional distribution based on the observed data. Several scenarios were considered. It was supposed that women who did not complete the questionnaires were in a poorer health state and possibly incurred higher costs. We assumed that the health state could have been 10% lower and costs would have been 10% higher than the missing at random setting. These assumptions were applied to two different settings, that (1) they were missing at the same value for both groups and (2) they were missing for one group only each time. In total, seven scenarios were considered.

Results

The aim of the following sections is to present the within-trial cost-effectiveness analysis.

Missing data

Data on intervention resource use and costs were missing for participants who withdrew consent, but their baseline data were still usable (five in the biofeedback PFMT group and two in the basic PFMT group). Details of missing participant-reported data on resource use in the primary and secondary care trial intervention are reported in *Table 17*. The proportions of missing data for cost and utility outcomes were equally distributed between the two groups. Complete total cost data were available for only 78 participants: 40 participants (13%) in the biofeedback PFMT group and 38 participants (13%) in the basic PFMT group. Utility data were available for 174 (biofeedback PFMT) and 158 (basic PFMT) participants. A summary of missing data for each cost variable, first year and total QALY is presented in *Table 17* (see *Appendices 13 and 14*).

As described in *Methods*, missing data are a significant issue for the calculation of complete-case cost and QALY pairs. This is a common issue in cost-effectiveness analyses alongside trials, in which data can be missing at an individual and item non-response level. It was therefore necessary to impute to address the significant missing data in the base-case cost-effectiveness analysis.

Resource use

Table 18 presents the resource use associated with the intervention and follow-up care in the biofeedback PFMT and basic PFMT groups.

Intervention

Overall, 264 out of 300 (88%) participants randomised to biofeedback PFMT received a home unit. Sixteen (44%) of those participants who did not get a biofeedback unit did not receive any intervention within the trial, five (14%) withdrew consent, seven (19%) withdrew after their first appointment, four (11%) received basic PFMT in error and the reason why the four (11%) remaining participants did not receive a home unit was unclear. One participant in the basic PFMT group was given a home biofeedback PFMT unit in error. Participants in both trial groups attended an average of four out of their six scheduled appointments, but the overall time spent at appointments was, on average, 26 minutes longer (95% CI 16.21 to 35.42 minutes) for those participants in the biofeedback PFMT group.

TABLE 17 Missing resource use/cost and QALY data

Variable	Treatment group, n missing/total N (%)	
	Biofeedback PFMT	Basic PFMT
Cost		
Intervention cost	5/300 (2)	2/300 (1)
Total GP appointments	201/300 (67)	200/300 (6)
Total nurse appointments	204/300 (68)	206/300 (69)
Total hospital doctor appointments	217/300 (72)	211/300 (70)
Total nurse/physiotherapist appointments	206/300 (69)	207/300 (69)
Total hospital stay	216/300 (72)	206/300 (69)
Total surgical interventions	219/300 (73)	214/300 (71)
Total medications	175/300 (58)	172/300 (57)
Total cost (intervention and follow-up)	260/300 (87)	262/300 (87)
Utility data (QALYs)	174/300 (58)	158/300 (53)

TABLE 18 Summary of intervention and follow-up care resource use (complete-case data)

		Treatment group						Difference ^a	95% CI
		Biofeedback PFMT (N = 300)			Basic PFMT (N = 300)				
Resource use	n	Mean	SD	n	Mean	SD			
Intervention									
Biofeedback units (number of home units)	295	0.89	0.31	298	0.00	0.06	0.90 ^b	0.86 to 0.95	
Number of scheduled trial appointments attended	295	4.19	1.88	298	4.05	2.09	0.23	−0.05 to 0.50	
Trial appointments' duration (minutes)	295	176.06	84.20	298	151.70	78.51	25.81 ^b	16.21 to 35.42	
Primary care (number of appointments)									
GP at the surgery	99	0.98	2.17	100	1.01	1.57	−0.10	−0.59 to 0.40	
Practice nurse at the surgery	96	0.34	1.22	94	0.14	0.45	0.17	−0.17 to 0.50	
Secondary care									
Hospital doctor	83	1.11	2.20	89	0.97	1.60	0.14	−0.42 to 0.70	
Hospital nurse/physiotherapist	94	2.02	2.88	93	1.96	3.59	0.33	−1.83 to .49	
Inpatient stay (number of nights) ^c	84	0.15	0.81	94	0.05	0.05	0.08	−0.01 to 0.17	
Surgical interventions for the treatment of SUI									
6 months	172	0.01	0.11	164	0.02	0.13	−0.01	−0.10 to 0.03	
12 months	204	0.04	0.19	210	0.05	0.22	−0.01	−0.06 to 0.04	
24 months	154	0.05	0.22	162	0.07	0.26	−0.02	−0.08 to 0.03	
Total	81	0.15	0.45	86	0.14	0.49	0.02	−0.09 to 0.13	
Prescribed medications									
6 months	209	0.18	0.46	204	0.19	0.46	−0.03	−0.14 to 0.08	
12 months	229	0.15	0.39	229	0.19	0.44	−0.05	−0.13 to 0.02	
24 months	168	0.20	0.46	174	0.17	0.41	0.02	−0.05 to 0.08	
Total	125	0.58	1.14	128	0.63	1.17	−0.10	−0.41 to 0.21	

a Differences have been adjusted for minimisation covariates and baseline EQ-5D-3L scores.

b Statistically significant.

c Non-surgery-related overnight hospital stay.

Follow-up care

Overall, the two trial groups had very similar levels of resource use in the follow-up period. Full details of all follow-up care resource use reported in the trial are provided in *Table 18*. The numbers of general practice doctor and nurse appointments were similar for both groups. The differences for both types of appointments were not statistically significant.

Although the biofeedback PFMT group had higher resource use for UI symptom management in secondary care (visits to hospital doctor, nurse or physiotherapist, and inpatient stay) than the basic PFMT group, the differences in resource use were not statistically significantly different. Furthermore, there was no significant difference in the number of surgical interventions (tension-free vaginal tape, bulking agents) or medications.

Intervention and follow-up costs

The differences in costs mirror the differences reported in resource use. Details of the cost results are reported in *Table 19*. As expected, the cost of biofeedback units was £33.76 (95% CI £31.99 to £35.52) higher in the biofeedback PFMT group. The cost of the appointments was also £53.12 (95% CI £34.20 to £72.05) higher in the biofeedback PFMT group. Both of these costs were statistically significantly higher. Therefore, the total cost of the within-trial intervention was significantly higher (£86.43, 95% CI £67.09 to £105.76) in the biofeedback PFMT group than in the basic PFMT group. The remaining cost differences were not statistically significantly different between the groups. Overall, the total cost was higher in the biofeedback PFMT group, but not statistically significant (difference was £120.97, 95% CI –£409.08 to £651.02). These findings are based on a complete-case analysis and it is notable that full-cost profile data were available for only 40 out of 300 (13%) participants and 38 out of 300 (13%) participants in the biofeedback PFMT and basic PFMT groups, respectively.

Generic and condition-specific quality-of-life measures

Results of EQ-5D-3L scores and adjusted differences between the groups are presented in *Table 20*. Participants randomised to the biofeedback PFMT group reported significantly lower EQ-5D-3L utility scores at baseline (mean difference biofeedback PFMT vs. basic PFMT –0.048, 95% CI –0.085 to –0.01). None of the follow-up scores was statistically significant between groups. The biofeedback PFMT group had lower QALYs (1.571) overall than the basic PFMT group (1.619) at 2 years, driven by the EQ-5D-3L imbalance in the baseline score. When adjusted for baseline differences, the incremental QALY (biofeedback PFMT vs. basic PFMT) was –0.041 (95% CI –0.121 to 0.039), which was not statistically significant.

TABLE 19 Summary of intervention and follow-up costs (complete-case data)

Resource use	Treatment group						Difference (£) ^a	95% CI (£)
	Biofeedback PFMT (N = 300)			Basic PFMT (N = 300)				
	n	Mean (£)	SD (£)	n	Mean (£)	SD (£)		
Biofeedback units	295	33.47	11.49	298	0.13	2.17	33.76 ^b	31.99 to 35.52
Within-trial appointments	259	353.89	169.23	298	304.92	157.80	53.12 ^b	34.20 to 72.05
Total intervention cost	295	387.36	175.62	298	305.05	157.88	86.43 ^b	67.09 to 105.76
Primary care								
GP at the surgery	99	36.25	80.16	100	37.37	57.97	−3.57	−21.90 to 14.76
Practice nurse at the surgery	96	3.73	13.25	94	1.50	4.93	0.17	−0.17 to 0.50
Secondary care								
Hospital doctor	83	130.43	242.20	89	117.12	189.60	12.81	−50.36 to 75.97
Hospital nurse/physiotherapist	94	166.81	237.82	93	160.90	291.97	27.82	−145.41 to 201.05
Inpatient stay	84	57.43	327.22	94	47.51	186.29	18.02	−8.81 to 44.86
Surgical interventions for the treatment of SUI	81	225.09	716.30	86	195.23	661.02	42.18	−103.07 to 187.43
Prescribed medications	125	49.95	113.15	128	58.11	165.32	−8.53	−56.25 to 39.19
Total overall cost	40	1261.42	1333.21	38	1118.26	1294.36	120.97	−409.08 to 651.02

^a Differences adjusted for baseline EQ-5D-3L score and minimisation covariates.

^b Statistically significant.

TABLE 20 Mean EQ-5D-3L scores at each time point, QALYs and differences between treatment groups

Variable	Treatment group						Difference ^a	95% CI
	Biofeedback PFMT (N = 300)			Basic PFMT (N = 300)				
	n	Mean	SD	n	Mean	SD		
EQ-5D-3L scores								
Baseline	287	0.788	0.288	295	0.836	0.240	−0.048 ^b	−0.085 to −0.010
6 months	225	0.785	0.300	229	0.823	0.256	−0.014	−0.055 to 0.026
12 months	235	0.769	0.320	241	0.811	0.267	−0.014	−0.056 to 0.028
24 months	166	0.810	0.284	171	0.808	0.264	0.003	−0.035 to 0.041
Visual analogue scale								
Baseline	280	73.74	19.39	285	73.90	19.65	1.10	−2.299 to 4.503
6 months	209	74.36	18.52	221	75.72	17.39	−1.68	−5.331 to 1.969
12 months	236	72.67	22.70	241	74.72	17.07	−2.09	−5.496 to 1.313
24 months	163	74.62	18.78	163	73.21	20.73	1.39	−3.171 to 5.951
QALYs								
First-year QALY	192	0.783	0.280	202	0.826	0.230	−0.02	−0.044 to 0.013
Second-year QALY ^c	126	1.571	0.492	142	1.619	0.455	−0.041	−0.121 to 0.039

^a All differences adjusted for minimisation covariates and baseline EQ-5D-3L score using ordinary least square regressions.

^b Significant at 5%.

^c Second year discounted at 3.5%.

The proportion of trial participants who reported problems in the domains of the EQ-5D-3L were equally distributed between the two groups (*Figure 5*). Pain/discomfort and anxiety/depression was the dimensions associated with most problems. Details of the number of participants are reported (see *Appendix 15*).

The mean ICIQ-LUTSqol scores at baseline and at each follow-up time point, as well as the differences between the two trial groups, are presented in *Table 21*. None of the differences in ICIQ-LUTSqol was statistically significant across the groups.

The proportion of participants in each trial group who reported any issues on each domain of the ICIQ-LUTSqol utility measure are reported in *Figure 6* (see *Appendix 16*). The problems reported by the highest proportion of participants were with physical activities, such as travelling, walking, running (between 79% and 97%), followed by emotional distress associated with their incontinence (between 60% and 83%) and limitations associated with their job or normal daily activities in and outside their household (between 53% and 84%), over the different time periods. The largest reduction in the proportion of participants reporting problems was observed between baseline and 6 months, and this occurred in both trial groups. The proportion of participants who reported issues remained stable for the rest of the follow-up (see *Figure 6*).

Costs incurred by and quality-adjusted life-years of cured and incontinent participants

At 24 months, the total cost of those participants who were reported as cured according to the primary clinical outcome (ICIQ-UI SF score) was £1032 (SD £942). The mean total cost for those who were still incontinent at the end of the follow-up period was £1203 (SD £1333). The mean total QALY for participants who were cured was 1.85 (SD 0.17), whereas it was 1.57 (SD 0.48) for those who were still incontinent. These results are based on only those who had complete data.

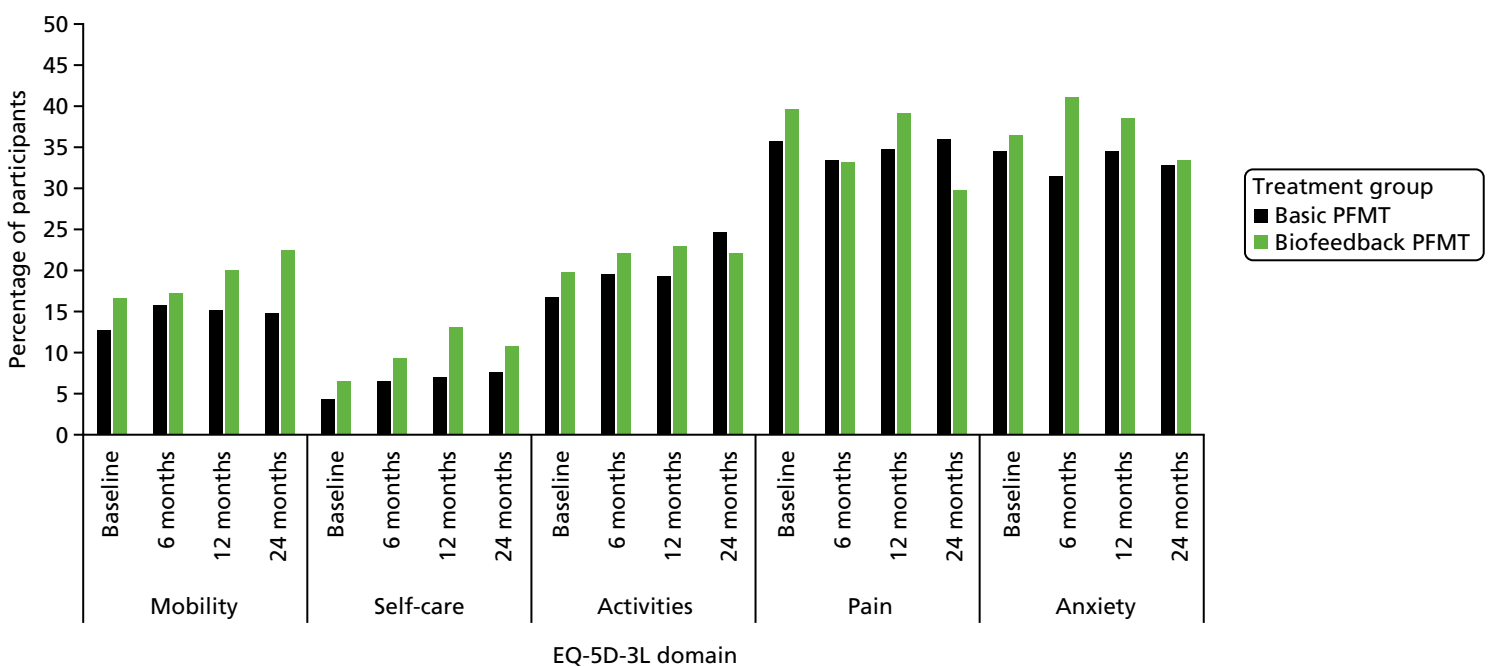


FIGURE 5 Proportion of participants experiencing any problems in each of the EQ-5D-3L domains.

TABLE 21 Mean ICIQ-LUTSqol scores at each time point, second-year QALYs and differences between treatment groups

Variable	Treatment group						Difference ^a	95% CI
	Biofeedback PFMT (N = 300)			Basic PFMT (N = 300)				
	n	Mean	SD	n	Mean	SD		
ICIQ-LUTSqol score at baseline	292	0.954	0.024	297	0.957	0.023	−0.003	−0.006 to 0.0001
ICIQ-LUTSqol score at 6 months	186	0.966	0.026	180	0.968	0.021	−0.003	−0.008 to 0.002
ICIQ-LUTSqol score at 12 months	189	0.967	0.026	184	0.969	0.022	−0.002	−0.007 to 0.003
ICIQ-LUTSqol score at 24 months	128	0.970	0.025	135	0.968	0.025	0.001	−0.006 to 0.008
First-year QALY	160	0.964	0.024	153	0.967	0.019	−0.000	−0.004 to 0.003
Second-year QALY (discounted)	98	1.900	0.046	100	1.903	0.038	−0.001	−0.011 to 0.008

a All differences are adjusted for baseline minimisation covariates and ICIQ-LUTSqol score (except the difference in baseline ICIQ-LUTSqol itself) using ordinary least square regressions.

Costs directly incurred by participants and carers and indirect costs of time and productivity lost due to urinary incontinence

There were no statistically significant differences in personal costs to the participants between the two groups (Table 22). The number of participants who reported the need to take full days off work because of incontinence was low ($n = 16$). Among those who required time off from usual activities in the follow-up period, the average number of days off in the basic PFMT group was 12.5 and the average number of days off in the biofeedback PFMT group was 10. The differences between groups were not statistically significant. The cost of self-purchased incontinence products (pads) constituted the highest annual cost to the participants: £806 for the biofeedback PFMT group and £862 for the basic PFMT group; however, the difference between the groups was not statistically significant. Only five participants reported self-purchasing other incontinence products, such as biofeedback units, vaginal cones or pessaries, in the follow-up period. The time and travel cost to the participant and carer (if applicable) of attending outpatient appointments was the second highest cost (on average, £200 for the biofeedback PFMT group and £303 for the basic PFMT group). Only very few participants (six in the basic PFMT group and eight in the biofeedback PFMT group) reported use of private secondary care (appointments to private doctors, nurses and physiotherapists). The total participant perspective cost difference (biofeedback PFMT vs. basic PFMT) was −£74.21 (95% CI −£201.21 to £52.78), but this was not statistically significant.

Base-case cost-effectiveness analysis

As described in *Missing data*, missing data were a significant issue for the calculation of complete-case cost and QALY pairs. This is a common issue in cost-effectiveness analyses alongside trials, in which data can be missing at an individual and item non-response level. Therefore, the remaining results (base-case and all sensitivity analyses) refer to analyses conducted on the multiply imputed data set, following best practice economic evaluation methodology. Results of the base-case analysis based on the reported intervention time indicated that, on average, the biofeedback PFMT group was £50 more costly than and had almost equivalent (+0.0009) QALYs to the basic PFMT group. The QALY differences are exceptionally small over a 2-year time horizon and equate to only 8 hours or 0.33 days improvement in quality of life over the trial time horizon. The base-case ICER was £56,617. The probability that biofeedback PFMT would be cost-effective was 48% and 49% at £20,000 and £30,000 WTP for a QALY thresholds, respectively (Table 23). The probability of being cost-effective remains unchanged at much higher WTP thresholds, as illustrated by the CEAC (Figure 7). The uncertainty is further illustrated by the scatterplots of incremental costs and incremental QALYs (Figure 8).

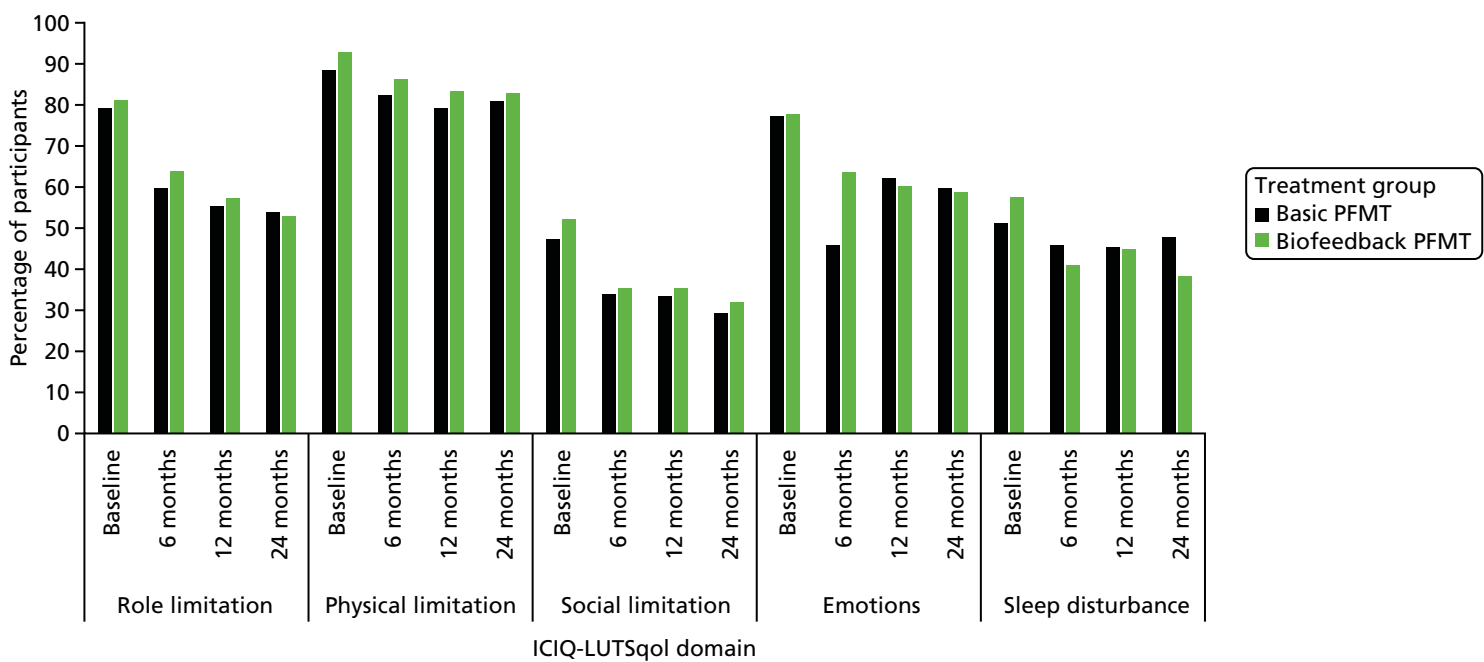


FIGURE 6 Proportion of participants experiencing any problems in each of the ICIQ-LUTSqol domains by group.

TABLE 22 Direct and indirect costs to participants and carers (complete-case data)

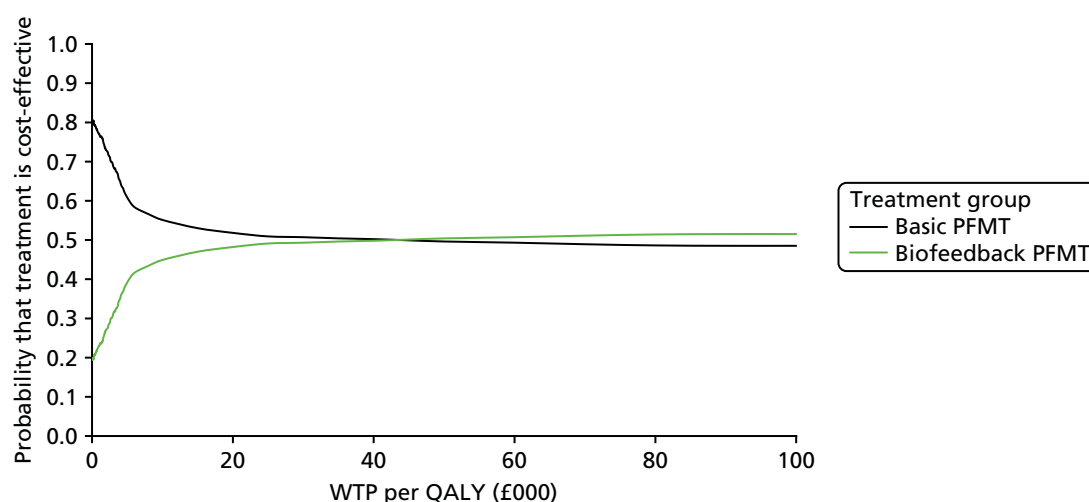
Cost	Treatment group							Difference (£) ^a	95% CI (£)
	Biofeedback PFMT (N = 300)			Basic PFMT (N = 300)					
	n	Mean (£)	SD (£)	n	Mean (£)	SD (£)			
Time/travel (inpatient stay)	66	52	216	78	19	80	29		–28 to 86
Time/travel (outpatient appointments)	73	200	196	73	303	539	–66		–163 to 32
Time and travel (general practice)	84	21	44	92	22	38	–4		–11 to 3
Self-purchased UI products	52	806	666	56	862	825	–127		–326 to 72
Private care	87	3	17	90	5	38	2 ^b		–3 to 6
Total cost	43	353	331	52	417	544	–74		–201 to 53

a Differences adjusted for baseline EQ-5D-3L and minimisation covariates.

b Adjusted differences using ordinary least squares regression due to GLM regressions not converging.

TABLE 23 Base-case cost-effectiveness results: therapist-reported duration of appointment (imputed data set)

Treatment group	Mean cost (£)	Difference in costs (£)	Mean QALYs	Difference in QALYs	ICER (£)	Probability (%) of being cost-effective	
						£20,000	£30,000
Basic PFMT	906		1.567			51.8	50.7
Biofeedback PFMT	956	50	1.568	0.0009	56,617	48.8	49.3

**FIGURE 7** Cost-effectiveness acceptability curves: therapist-reported duration of appointment (imputed data set).

Sensitivity analysis

Recommended appointment time was based on the assumptions that there was no difference in the duration of the appointments and that both biofeedback PFMT and basic PFMT appointments were similar (1 hour for the first appointment and 30 minutes for each subsequent appointment). The results indicate that, on average, the biofeedback PFMT group's costs were £23 higher than the costs of basic PFMT group, but the biofeedback PFMT group had 0.004 fewer QALYs than the basic PFMT group (Table 24). Biofeedback PFMT group was dominated by the basic PFMT group, as it cost more and was less effective. The probability that biofeedback PFMT was cost-effective was 42% at the £20,000 WTP threshold and 43% at the £30,000 WTP threshold (Figures 9 and 10).

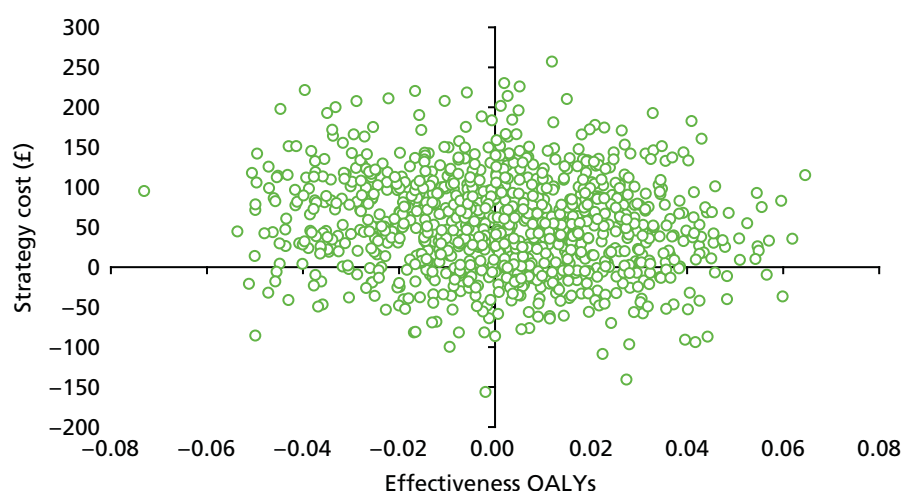


FIGURE 8 Scatterplot of incremental costs and QALYs for biofeedback PFMT vs. basic PFMT: therapist-reported duration of appointment (imputed data set).

TABLE 24 Recommended appointment time (3.5 hours for six appointments) (imputed data set)

Treatment group	Mean cost (£)	Difference in costs (£)	Mean QALYs	Difference in QALYs	ICER (£)	Probability (%) of being cost-effective	
						£20,000	£30,000
Basic PFMT	885		1.570			58.2	57.4
Biofeedback PFMT	909	23	1.566	-0.004	Dominated	41.8	42.6

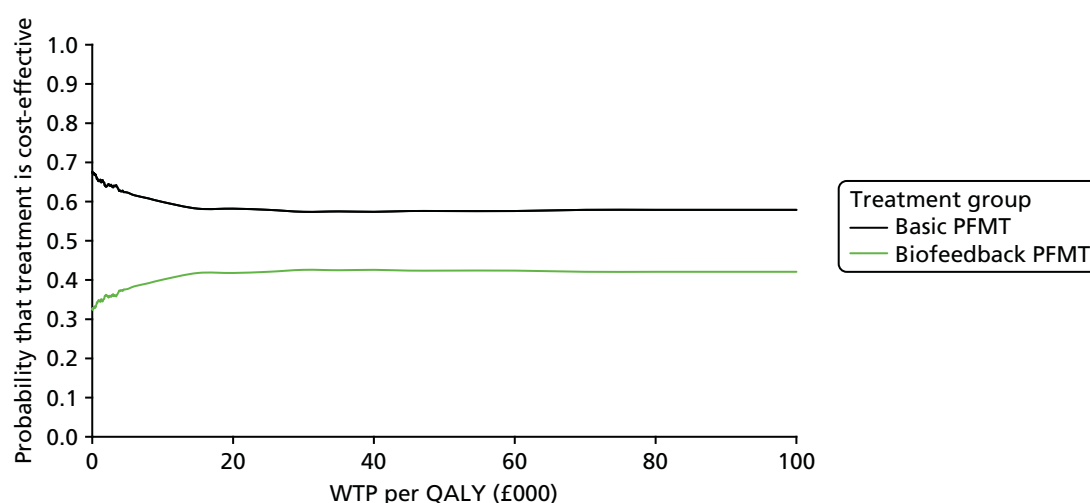


FIGURE 9 Cost-effectiveness acceptability curves: recommended appointment time (3.5 hours for six appointments).

Subgroup analysis

The results of the sensitivity analyses exploring the cost-effectiveness of biofeedback PFMT compared with basic PFMT for different subgroups are detailed in *Table 25*. The results of those participants with SUI were similar to those of the base-case analysis (for sensitivity analyses CEACs and ICER scatterplots, see *Appendix 17, Figures 13–32*).

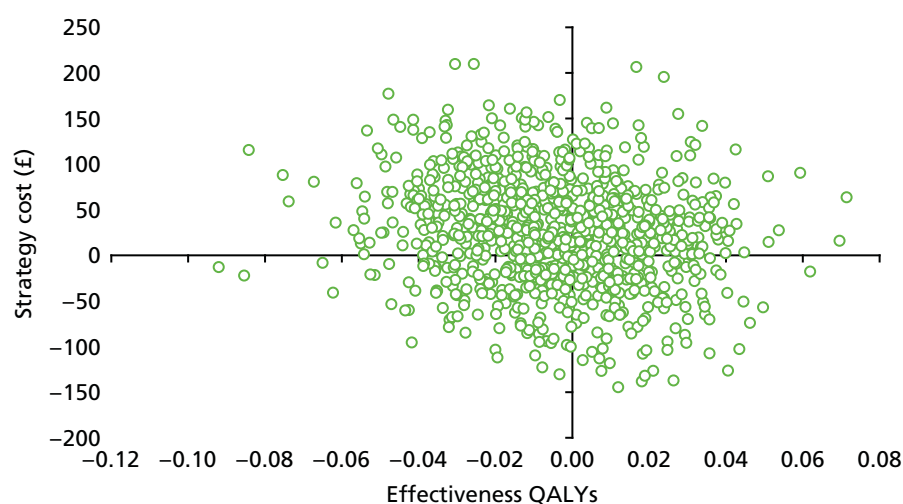


FIGURE 10 Scatterplot of incremental costs and QALYs for biofeedback PFMT vs. basic PFMT: recommended appointment time (3.5 hours for six appointments).

TABLE 25 Sensitivity analyses

Subgroup	Cost (£)	Incremental cost (£)	QALYs	Incremental QALYs	ICER (£)	Probability (%) of being cost-effective	
						£20,000	£30,000
SUI (n = 235)							
Basic PFMT	815		1.596			61.3	56.2
Biofeedback PFMT	976	162	1.599	0.003	57,936	38.7	43.8
MUI (stress and urgency) (n = 365)							
Biofeedback PFMT	939		1.551			53.2	52.1
Basic PFMT	993	54	1.550	−0.001	Dominated	46.8	47.9
Aged < 50 years (n = 360)							
Basic PFMT	890		1.573			88.8	88.4
Biofeedback PFMT	919	29	1.540	−0.034	Dominated	11.2	11.6
Aged ≥ 50 years (n = 240)							
Basic PFMT	951		1.560			8.0	7.1
Biofeedback PFMT	1033	82	1.618	0.058	1417	92.0	92.9
Lower severity level (ICIQ-UI SF score of < 13) (n = 250)							
Basic PFMT	827		1.603			19.4	20.1
Biofeedback PFMT	863	36	1.629	0.026	1397	80.6	79.9
Higher severity level (ICIQ-UI SF score of ≥ 13) (n = 350)							
Basic PFMT	977		1.543			68.2	67.8
Biofeedback PFMT	1036	59	1.529	−0.014	Dominated	31.8	32.2
Using ICIQ-LUTSqOL scores as a utility measure							
Basic PFMT	906		1.896			46.2	42.9
Biofeedback PFMT	956	50	1.896	0.000	177,668	53.8	57.1
Undiscounted costs and QALYs							
Basic PFMT	897		1.595			54.3%	53.4%
Biofeedback PFMT	948	51	1.594	−0.001	Dominated	45.7%	46.6%

TABLE 25 Sensitivity analyses (continued)

Subgroup	Cost (£)	Incremental cost (£)	QALYs	Incremental QALYs	ICER (£)	Probability (%) of being cost-effective	
						£20,000	£30,000
Discounted costs and QALYs at 6%							
Basic PFMT	886.07		1.550			55.3%	54.1%
Biofeedback PFMT	938.22	52.16	1.549	−0.001	Dominated	44.7%	45.9%
Unadjusted costs and QALYs							
Basic PFMT	904.36		1.602			96.9	96.9
Biofeedback PFMT	953.31	48.95	1.532	−0.070	Dominated	3.1	3.1
Same MNAR parameters in two groups							
−10% quality of life in both groups							
Basic PFMT	906		1.508			55.9	54.9
Biofeedback PFMT	956	50	1.507	−0.001	Dominated	44.1	45.1
+10% cost in both arms							
Basic PFMT	980		1.567			51.6	50.0
Biofeedback PFMT	1033	55	1.568	0.001	61,855	48.4	50.0
−10% quality of life and + 10% cost							
Basic PFMT	980		1.508			56.3	54.7
Biofeedback PFMT	1033	55	1.507	−0.001	Dominated	43.7	45.3
Different MNAR in the two groups							
−10% quality of life in biofeedback PFMT group							
Basic PFMT	906		1.567			99.6	99.6
Biofeedback PFMT	956	50	1.507	−0.060	Dominated	0.4	0.4
−10% quality of life in basic PFMT group							
Basic PFMT	906		1.507			0.8	0.7
Biofeedback PFMT	956	50	1.568	0.060	843	99.2	99.3
+10% cost in biofeedback group							
Basic PFMT	906		1.567			57.9	55.2
Biofeedback PFMT	1033	124	1.568	0.001	140,278	42.1	44.8
+10% cost in basic PFMT group							
Biofeedback PFMT	956		1.568			45.3	45.4
Basic PFMT	980	−19	1.567	0.001	Dominated	54.7	54.6

Type of incontinence

On average, the biofeedback PFMT intervention cost £152 more and was more effective than basic PFMT. The ICER was £57,936 and the probability that biofeedback PFMT would be considered to be cost-effective was 39% at society's £20,000 WTP threshold and 44% at the £30,000 threshold. The results of those participants with MUI differed from the base-case results. On average, the biofeedback PFMT intervention cost less and was more effective than basic PFMT. Biofeedback PFMT dominated basic PFMT. The probability that biofeedback PFMT would be considered to be cost-effective was 53% at society's £20,000 WTP threshold and 52% at the £30,000 threshold.

Age

The results of those participants aged < 50 years were also different from the base-case analysis. On average, the biofeedback PFMT intervention cost more and was less effective than basic PFMT. Biofeedback PFMT was dominated by basic PFMT. The probability that biofeedback PFMT would be considered to be cost-effective was 11% at society's £20,000 WTP threshold and 12% at the £30,000 threshold. For participants aged ≥ 50 years, biofeedback PFMT cost, on average, £82 more than basic PFMT, and the difference in QALYs was 0.058 (vs. 0.0009 in the base-case analysis). The ICER was £1417. The probability that biofeedback PFMT would be considered to be cost-effective was 92% at society's £20,000 WTP threshold and 93% at the £30,000 threshold.

Severity of incontinence

The results for those participants with a lower-severity ICIQ-UI SF score (< 13) indicated that, on average, biofeedback PFMT cost £29 more and was 0.026 QALYs more effective than basic PFMT. The ICER was £1397. The probability that biofeedback PFMT would be considered to be cost-effective was 81% at society's £20,000 WTP threshold and 80% at the £30,000 threshold. The results for those participants with a higher-severity ICIQ-UI SF score (≥ 13) indicated that, on average, biofeedback PFMT cost £59 more and was 0.014 QALYs less effective than basic PFMT. The biofeedback PFMT group was dominated by the basic PFMT group. The probability that biofeedback PFMT would be considered to be cost-effective was 11% society's £20,000 WTP threshold and 12% at the £30,000 threshold.

Quality-of-life scores

The results of the analysis using the ICIQ-LUTSqol were similar to those of the base-case analysis. On average, biofeedback PFMT cost £50 more and was more effective than basic PFMT. The ICER was £177,668 and the probability that biofeedback PFMT would be considered to be cost-effective was 54% at society's £20,000 WTP threshold and 57% at the £30,000 threshold.

Discounting and unadjusted costs and quality-adjusted life-years

The results of the discounting and unadjusted analyses differ from the base-case analysis. On average, biofeedback PFMT cost more and was less effective than basic PFMT; therefore, biofeedback PFMT was dominated, although the differences were not statistically significant. The change in results is mainly because the baseline imbalance was maintained, as the data were not adjusted to take into account this difference.

Assuming that data were missing not at random

As shown in *Table 25* (also see *Appendix 17*), this assumption has an impact on the results, in which the missing quality-of-life data are 10% lower in the biofeedback PFMT group only and the QALY difference is statistically significantly lower -0.060 (95% CI -0.117 to -0.003) for the biofeedback PFMT group. Biofeedback PFMT is dominated by basic PFMT as it costs more and is less effective, and the probability that biofeedback PFMT would be considered to be cost-effective was 0.4% at society's £20,000 WTP for a QALY threshold and 0.4% at the £30,000 WTP threshold. The results are reversed when we assume a reduction of 10% in the missing of quality-of-life data in the basic PFMT group only, as the QALY is statistically significantly higher (0.060, 95% CI -0.04 to -0.115) for the biofeedback PFMT group. The ICER for the biofeedback PFMT group is £843, as it costs more and is more effective, and the probability that biofeedback PFMT would be considered to be cost-effective was 99% at society's £20,000 WTP for a QALY threshold and 99% at the £30,000 WTP threshold. The rest of the scenario results were the same as those reported when it was assumed that data were missing at random.

Discussion

The base-case results indicate that biofeedback PFMT, on average, cost more than basic PFMT and generated more QALYs. However, none of these results was statistically significant. Although, on average, the total number of appointments attended was similar in both groups, the total duration of the biofeedback PFMT

appointments was 26 minutes longer than the basic PFMT appointment. The difference in the intervention cost was driven by the cost of the biofeedback PFMT units and the duration of the appointments. The difference in appointment duration was attributed to information technology (IT) problems encountered by therapists during appointments and the extra time associated with setting up the biofeedback PFMT equipment. However, over the 24-month follow-up period, this cost difference was reduced, as there were no statistically significant differences in primary and secondary care resource use and costs between the two groups.

At baseline, the EQ-5D-3L score was statistically significantly higher for the basic PFMT group than for the biofeedback PFMT group; therefore, all the analyses conducted were adjusted for the baseline EQ-5D-3L utility scores. This difference decreased and was not statistically significant at any time point during the follow-up period. The adjusted mean QALY difference (biofeedback PFMT vs. basic PFMT) was negligible, at 0.0009. Over the 2-year follow-up period, this would equate to 8 hours' improvement in quality of life, which may not be meaningful. The ICER for analysis based on the reported duration of appointments was £56,617, which is higher than the £20,000–30,000 WTP threshold typically used in the UK.³⁹ Therefore, biofeedback PFMT would not be considered to be cost-effective. The probability that biofeedback PFMT was cost-effective if society were willing to pay £20,000 for an additional QALY was 48% and did not increase above 55% at higher WTP thresholds. However, some subgroup analyses led to changes in the base-case cost-effectiveness conclusions. This could be because the trial was not powered to undertake these analyses.

Deterministic sensitivity analyses were conducted to address areas of uncertainty in assumptions and data collection. The conclusions of the cost-effectiveness analysis that addressed the uncertainty in the cost and QALY data were similar to those of the base case. The first analysis assumed that there were no differences in the duration of the appointments. As expected, this analysis reduced the magnitude of incremental costs, but the resultant ICER remained highly uncertain and the probability of biofeedback PFMT being cost-effective increased by only 1%. This uncertainty is depicted by the distribution of the simulations in all quadrants of the cost-effectiveness plane.

The number of participants who reported full days off work as a result of incontinence was low. Among those who required time off their usual activities in the follow-up period, the average number of days off in the basic PFMT group was 12.5 and in the biofeedback PFMT group was 10. The differences were not statistically significant.

The amount of uncertainty around the final cost and QALY increments was found to vary, with point estimates being highly unstable and changing sign and magnitude with every attempt at multiple imputation using chained equations. This shows that the two interventions for UI are very similar in terms of cost-effectiveness. As indicated in the base-case analysis, the main driver of the costs was the biofeedback units, as the subsequent costs were, on average, similar in the two groups. Thus, there was an additional cost associated with the biofeedback PFMT intervention but no difference in QALYs.

One of the challenges for the cost-effectiveness analysis was the number of missing data. The analysis of complete-case QALY data at 12 and 24 months ignored important individual-level information for participants who reported no or partially complete EQ-5D-3L data at any time point. However, missing data were evenly distributed between the two trial groups at all time points, with 98 out of 300 (33%) and 108 out of 300 (36%) having missing 12-month QALY data in the basic PFMT and the biofeedback PFMT group, respectively, and this increased to > 50% in both groups: 53% and 58%, respectively. This challenge was addressed by imputing cost and QALY outcomes. The base-case analysis was based on multiple imputation. The conclusions of the imputed analysis were similar to those of complete-case analysis. Analyses assuming a 10% reduction in the missing quality-of-life data led to changes in the overall results, as both treatments had higher probability of being cost-effective when the other group had the reduction attached to their values. This can be attributed to the fact that there was a very small difference in the QALYs in the base-case analysis (0.0009) and a reduction in one group could favour the

other group. Analyses that assume that cost data were MNAR did not change the overall conclusions, as the QALY difference remained low, and, taking into account the results from the clinical effectiveness analysis, we do not have any reason to believe that the pattern of missing cost and QALY data was significantly different across treatment arms, or can be explained by factors for which we are unable to control in our regression models.

One of the key strengths of the cost-effectiveness analysis is that it was conducted alongside a multicentre trial, adding to the generalisability of the results. A detailed costing of the intervention was undertaken and participants were followed up over 24 months. The quality-of-life data were collected using both generic and condition-specific instruments at several time points to ensure that the impact on quality of life was captured. As the trial was not powered to detect differences in economic outcomes, conclusions from the economic evaluation have to be based on the consideration of the balance of probabilities.

Conclusion

The base-case analysis showed that biofeedback PFMT was not significantly more expensive than basic PFMT and did not generate significantly higher QALYs. The ICER (£56,617) is greater than society's WTP threshold of £30,000. Biofeedback PFMT was associated with a 49% chance of being cost-effective if society is willing to pay £30,000 for a QALY. The difference in intervention costs was driven by the cost of the biofeedback units and longer duration of biofeedback PFMT appointments. A sensitivity analysis assuming that the duration of appointments was the same for both groups indicated a reduction in the cost difference. Biofeedback PFMT was dominated by basic PFMT, as it cost more but was less effective. The chance that biofeedback PFMT would be considered cost-effective at the £30,000 threshold for society's WTP for a QALY reduced to 48%. These results have to be interpreted taking into account the uncertainty surrounding the estimates of cost and QALYs.

Chapter 5 Process evaluation

Introduction

This chapter summarises the OPAL trial process evaluation, addressing the third research aim (see *Chapter 1, Aims and objectives*), and is in line with guidance for evaluating complex interventions.^{16,53} The two OPAL trial interventions are described in *Chapter 2* and their development was underpinned by the IMB model,¹⁷ the BCT taxonomy⁸ and self-efficacy theory⁵⁴ (see also *Chapter 1* and the trial protocols^{20,21}).

The aim was to:

- identify and investigate the possible mediating factors that affect the effectiveness of the intervention (including fidelity to intervention delivery and uptake), how these mediating factors influence effectiveness and whether or not the factors differ between randomised groups.

The process evaluation was conducted in parallel with the main trial and comprised analysis of multiple data sets generated throughout the trial, including:

1. therapist-completed TAF checklists
2. audio-recorded appointments
3. interviews with therapists
4. PFMT exercise diaries
5. patient-completed trial baseline and follow-up questionnaires.

This chapter presents the research methods and results from the first four data sources, reporting findings from the intervention period.

Methods

This mixed-methods process evaluation had a concurrent design. For intervention delivery, we assessed the 'fidelity of form'⁵⁵ for the core components of the two interventions (i.e. the core components that were to be used for all participants and not individualised). We also assessed 'fidelity of function'⁵⁵ for individualised or optional components, as therapists were given the intervention protocols to follow, with the expectation that they would tailor some components to individual participants (in terms of the exact exercise prescription, lifestyle advice, etc.). The main data sources for assessing intervention delivery were therapist-completed TAF checklists, audio-recordings of appointments and therapist interviews.

Although participants' fidelity to the intervention, with respect to their experiences of receiving treatment in clinic and how they engaged with PFMT at home, was mainly covered by the longitudinal qualitative case study (see *Chapter 6*), we also evaluated this using the audio-recording of appointments, therapist interviews and the free-text responses in participants' exercise diaries.

Process evaluation sampling, recruitment strategy and data collection

Appointment checklists

A protocolised checklist (basic PFMT group or biofeedback PFMT group) was completed by all therapists for each appointment (for examples of core and optional components for each group, see *Appendix 18*). These checklists were embedded in the TAF completed at each appointment [see the project web page: www.journalslibrary.nihr.ac.uk/programmes/hta/117103/#/ (accessed 29 July 2019)]. Each checklist contained the core and optional intervention components for the appointment, and varied by appointment

number (1–6) and by group. For each appointment, the biofeedback PFMT checklists had all the core basic PFMT components and additional biofeedback-related components. Therapists ticked yes or no for each component to allow assessment of fidelity. At the end of the tick-box section, an open-text space enabled therapists to explain the clinical reasoning for any omissions or additions to tailor the intervention based on a participant's needs.

Appointment audio-recordings

Audio-recordings were planned to obtain a deliberately heterogeneous purposive sample of 100 appointments from across both basic PFMT and biofeedback PFMT groups; from the different appointment time point (1–6); centres and UI type and severity; and different therapist type (physiotherapist/nurse). Challenges with arranging recordings and the increase in the number of centres meant that we subsequently focused sampling on the first and last appointments, as a result of our a priori hypothesis that treatment delivery may be more intensive and concentrated in these appointments. The first appointment was scheduled to be longer (1 hour) and involved assessment (including assessment of correct contraction), education and teaching for participants on how to properly perform PFMEs, use of clinic biofeedback and teaching use of home biofeedback (for the biofeedback PFMT group only) and explanation of any recommended changes to lifestyle. However, as occurs in clinical practice, some of these elements could be deferred to the second appointment based on the individual needs of the woman. The last appointment (30 minutes scheduled) was important because supervised therapy was ending and participants were being given important information to allow them to self-care, such as instructions regarding the maintenance PFMT dose. Appointments 2–5 were also 30 minutes in duration and component content was mostly similar (i.e. repeated) across these four appointments.

If participants had signed on their consent form that they were willing to have an appointment audio-recorded, the researcher purposively selected participants and telephoned them to ask if they were willing to have a specific appointment recorded, answer any questions they had and obtain verbal consent. Therapists were then informed which participants and appointments to record, using a centre-specific password-protected digital audio-recorder; these were returned to the researcher at regular intervals for downloading data. Recordings were anonymised and transcribed verbatim.

Therapist interviews

Semistructured telephone interviews were undertaken with staff at the end of their centre's participation in the delivery of the trial interventions. Initially, we planned to interview at least one therapist from each centre; however, for various reasons this was not possible and a decision was taken by the process evaluation team to widen recruitment to include at least one person from each centre, but extending the sample to administrative support staff and nurses involved only in recruiting women. In the first instance, the interview researcher e-mailed the therapist(s) at each centre, informing them about the interview purpose and inviting them to participate. Their reply, with a suitable date and time, was accepted as consent to be interviewed. A topic guide was developed from the literature, from issues arising during delivery of the trial interventions and from case study participant interviews. The resultant broad categories of questions were delivery of the interventions to participants; response of participants; maintenance of PFMT; and aspects of theory underlying the intervention. In these categories, the interviewer deliberately sought therapist perspectives on their protocol adherence and participants' PFMT adherence. Interviews were audio-recorded, anonymised and transcribed verbatim then uploaded to NVivo 11 (QSR International, Warrington UK) for coding.

PFMT exercise diaries

Exercise diaries were provided to each participant at appointments 1–5, to be completed as a record of PFMT undertaken at home and returned at the following appointment. Therapists wrote the agreed PFMT (and biofeedback use) to be followed (the 'dose') in the diary and participants were asked to sign their agreement (a BCT called 'commitment') to undertake home PFMT and, if allocated, home biofeedback PFMT. At the following appointment, exercise diaries were collected, progress reviewed and a new diary issued. Diaries provided space for participants to record, on a daily basis, the number of sets of exercises undertaken and, if allocated, their use of biofeedback PFMT, along with any free-text comments about their homework.

Process evaluation analysis

In this chapter, the analysis of each individual data source, undertaken to reach separate conclusions, is presented.

Appointment checklists

The number of available checklists (i.e. a checklist for all appointments attended) for participants in both groups was compared with the potential number. In each appointment, the number of components checked 'yes' were summarised descriptively (means, modes and interquartile ranges) to report the extent to which therapists delivered the core and optional components of each appointment in each group. Therapist free-text comments about intervention omissions, additions and tailoring were coded using a framework developed following content analysis of a 10% representative sample of appointments.

Appointment audio-recordings

A quantitative coding scheme was developed based on the appointment checklists. The coding scheme for each appointment contained explicit guidance for assessing whether or not the protocol components (core and optional) were delivered (yes/no), followed by an assessment of the quality of delivery. The reliability of the coding scheme was checked by comparing the codes independently assigned by three members of the research team to a sample of audio-recordings. Following discussion of differences in coding, the agreed coding scheme was finalised. Using both the audio-recordings and the transcribed audio-recordings, the coding was completed and entered into IBM SPSS Statistics 25 (IBM Corporation, Armonk, NY, USA) for data management. Coded data were summarised descriptively by appointment and by group.

Therapist interviews

Using a general inductive and descriptive qualitative approach, analysis began with familiarisation with the data. A thematic framework⁵⁶ was developed by one researcher from initial coding of 10 interviews; a second researcher then independently coded four transcripts, and a third researcher cross-checked and agreed the final set of codes. This coding framework was then applied across the data set. Findings are reported in alignment with the sequence of issues discussed in the interviews, with a focus on the therapists' perspectives on the barriers to and facilitators of intervention fidelity, and participants' engagement and adherence, to demonstrate variation over time.

PFMT exercise diaries

Content analysis was used to summarise free-text entries made by participants, checking for similarities and differences between groups and appointments. This analysis took place prior to the main trial results being known. Subsequent to knowing the results, we examined the returned diaries for completion of the therapist- and participant-signed PFMT agreement. This BCT was a core component in both trial interventions and was summarised by appointment and by group as a proportion of those attending each appointment.

Results

We present results according to the four main process evaluation data sources.

Appointment checklists

The potential maximum number of checklists was 3600 (600 participants, six appointments each). The number of available checklists by group and by appointment is summarised in *Table 26*, and does not necessarily reflect how well the checklists were completed by the therapist. Missing checklists include participants who withdrew, participants who did not attend a specific appointment, non-submission of a checklist by a therapist (some forgot to complete it) or a missing TAF. There were no clear differences between groups in the proportion of available checklists, either by specific appointment or over time. There was a steadily decreasing trend from appointment 1 (91% submitted) to 6 (60% submitted).

TABLE 26 Number of available checklists compared with potential total number of checklists by group and by appointment

Appointment number	Treatment group, n (%)		Total (N = 600), n (%)
	Basic PFMT (N = 300)	Biofeedback PFMT (N = 300)	
1	265 (83)	279 (93)	544 (91)
2	235 (78)	245 (81)	480 (80)
3	189 (63)	198 (66)	387 (65)
4	175 (58)	185 (62)	360 (60)
5	166 (53)	156 (52)	322 (54)
6	175 (58)	182 (61)	357 (60)
Total	1205 (67)	1245 (69)	2450 (68)

On the checklists, therapists ticked 'yes' or 'no' to delivery of each core component for each appointment [for examples, see the project web page: www.journalslibrary.nihr.ac.uk/programmes/hta/117103/#/ (accessed 29 July 2019)]. *Table 27* summarises the number of core components ticked 'yes', by group and by appointment. Between the two groups, there were no observed differences in the number of basic core components (i.e. components that were the same in both groups and ticked 'yes' by therapists per appointment). Therapists most often ticked either a full complement of core components or just one less than a full complement; this was the case for every appointment apart from appointment 5 in the biofeedback PFMT group, when the mode was only 15 of a possible 18 ticks. The interquartile ranges reveal a skewed data distribution in favour of a high number of basic PFMT core components being ticked 'yes' for all appointments and for both groups (see *Table 27*, columns 2 and 3).

In the biofeedback PFMT group only, therapists most often ticked 'yes' to either all the biofeedback core components or one less than a full complement. In all but the last biofeedback PFMT appointment, at least 75% of the checklists had more than half of the core components ticked 'yes' (see *Table 27*, column 4). When combining the basic and biofeedback core components (see *Table 27*, column 5), it can be seen that the biofeedback PFMT group received more 'yes' ticks overall than the basic PFMT group. For example, in appointment 1, the median number of ticks was 18 for the basic PFMT group (out of a possible 19, see *Table 27*, column 2) and for the biofeedback PFMT group, the median was 26 (out of a possible 28, see *Table 27*, column 5). This pattern was similar for all appointments, indicating that, based on therapist reporting, the biofeedback PFMT group did have more core components delivered during appointments, consistent with the intervention protocol.

A similar summary for the optional intervention components [see the project web page: www.journalslibrary.nihr.ac.uk/programmes/hta/117103/#/ (accessed 29 July 2019)] revealed a very different pattern: very few were ticked 'yes' in any appointment, for either group (*Table 28*). The mode is consistently zero for all appointments in both groups; the interquartile range is much larger and in only appointment 6 does the median reach half the number of available basic optional components. The optional component pattern is consistent across the appointments and between the groups for the basic PFMT components. The biofeedback PFMT group did receive some additional biofeedback-specific optional components, but relatively few compared with those available on the checklist (see *Table 28*, column 3).

Appointment audio-recordings

As described in *Methods*, collecting the full intended sample of audio-recordings was challenging because therapists were unfamiliar with equipment, declined to do their quota because of staff shortages or forgot to record or identify the patient/appointment number, or because participants failed to attend the selected appointment or withdrew from the trial. *Table 29* demonstrates our decision to focus on collecting data from appointments 1–6, as appointments 2–5 were designed to repeat intervention content from previous

TABLE 27 Summary of checklist core components ticked for each appointment by group: differentiating between basic PFMT core components and biofeedback-specific PFMT core components

Appointment	Treatment group			
	Basic PFMT		Biofeedback PFMT	
	Basic core components	Basic core components	Biofeedback core components	Basic and biofeedback core components
Appointment 1				
Potential range	0–19	0–19	0–9	0–28
<i>n</i>	265	279	279	279
Median	18	18	8	26
Mode	18	18	9	27
IQR	17–19	15–18	5–9	20–27
Appointment 2				
Potential range	0–20	0–20	0–13	0–33
<i>n</i>	235	245	245	245
Median	18	18	11	29
Mode	19	19	12	31
IQR	17–19	16–19	9–12	25–31
Appointment 3				
Potential range	0–18	0–18	0–14	0–32
<i>n</i>	189	198	198	198
Median	16	16	12	28
Mode	17	17	13	30
IQR	15–17	13–17	10–13	24–30
Appointment 4				
Potential range	0–19	0–19	0–13	0–32
<i>n</i>	175	185	185	185
Median	17	16	11	27
Mode	18	18	12	30
IQR	14–18	14–18	9–12	22–30
Appointment 5				
Potential range	0–17	0–17	0–13	0–30
<i>n</i>	166	156	156	156
Median	15	14	12	26
Mode	17	16	13	26
IQR	12–16	12–16	10–13	21–28
Appointment 6				
Potential range	0–19	0–19	0–8	0–27
<i>n</i>	175	182	182	182
Median	18	17	6	22
Mode	18	18	6	25
IQR	15–18	13–18	3–6	17–24
IQR, interquartile range.				
Notes				
<i>n</i> = number of available checklists.				
Remaining data summarise number of components ticked 'yes'.				

TABLE 28 Summary of checklist optional components ticked for each appointment by group: differentiating between basic PFMT optional components and biofeedback-specific PFMT optional components

Appointment	Treatment group			
	Basic PFMT	Biofeedback PFMT		
	Basic optional components	Basic optional components	Biofeedback optional components	Basic and biofeedback optional components
Appointment 1				
Potential range	0–11	0–11	0–6	0–17
<i>n</i>	265	279	279	279
Median	3	3	3	6
Mode	0	0	0	0
IQR	0–7	0–6	0–4	0–10
Appointment 2				
Potential range	0–18	0–17 (1 omitted in error)	0–12	0–29
<i>n</i>	235	245	245	245
Median	8	5	5	10
Mode	0	0	0	0
IQR	4–12	1–9	0–8	2–17
Appointment 3				
Potential range	0–15	0–14 (1 omitted in error)	0–12	0–26
<i>n</i>	189	198	198	198
Median	6	4	4	9
Mode	0	0	0	0
IQR	3–10	0–8	0–8	0–16
Appointment 4				
Potential range	0–13	0–13	0–8	0–21
<i>n</i>	175	185	185	185
Median	6	3	2	6
Mode	0	0	0	0
IQR	2–9	0–7	0–5	2–12
Appointment 5				
Potential range	0–15	0–13 (2 omitted in error)	0–9	0–22
<i>n</i>	166	156	156	156
Median	6	3	2	6
Mode	0	0	0	0
IQR	2–10	0–8	0–4	0–11
Appointment 6				
Potential range	0–16	0–12 (4 omitted in error)	0–10	0–22
<i>n</i>	175	182	182	182
Median	8	6	1	7
Mode	0	0	0	0
IQR	4–12	2–9	0–3	3–11

IQR, interquartile range.

Notes*n* = number of available checklists.

Remaining data summarise number of components ticked 'yes'.

TABLE 29 Summary of audio-recordings made by appointment and by group

Treatment group	Appointment (n)						Total (n)
	1	2	3	4	5	6	
Basic PFMT	7	6	8	6	8	11	46
Biofeedback PFMT	8	6	5	6	5	12	42
Total	15	12	13	12	13	23	88

appointments, and, although appointment 2 in the biofeedback group could be when the device was first introduced, it was not feasible to know this in advance when arranging the appointments to be audio-recorded. Although we achieved more recordings in appointments 1 and 6, recruiting to appointment 1 remained difficult, but we achieved a reasonably balanced sample across groups (see *Table 29*). Overall, the number of audio-recordings made was 88% of the original target.

The number of basic PFMT and biofeedback PFMT core components that were heard being delivered are summarised in *Table 30*. We agreed that some components were unlikely to be heard and these were not coded; hence, the potential range of components differs from the checklist summaries.

Therapists were heard to use fewer core components, overall, than the number that were potentially audible for coding. This was the case for both groups and across all appointments. The biofeedback PFMT group received more core components (basic and biofeedback combined, see *Table 30*, column 5) than the basic PFMT group. No participant in the biofeedback PFMT group started biofeedback for the first time in session 2.

A similar summary of audio-recording data was not undertaken for the optional intervention components, given that we knew from the checklist data that therapists were not reporting the use of many of these optional components.

Therapist interviews

We were able to recruit from 21 of 23 centres, obtaining 30 interviews from 26 physiotherapists, one nurse continence specialist (total of 27 therapists delivering the interventions), two nurses and one administrator involved in a variety of tasks, including participant recruitment, obtaining consent and dealing with IT issues.

We report therapist perceptions of recruitment and delivering the trial overall, their perceptions of both basic and biofeedback PFMT interventions and any differences between them. We pay particular attention to therapists' perspectives on (1) their protocol fidelity and (2) participants' intervention adherence.

Recruitment of trial centres and participants

Therapists said that the financial incentive was important during recruitment of their centre to the trial and for supporting them in continuing in the trial. Management were more likely to agree if there was funding associated with participation and the larger English centres with research experience knew this enabled management buy-in:

I think that was, that was important, if it was just me doing extra work and the department not making money out of it, they probably wouldn't have been as supportive, 'cause they rely on that money kinda coming in for different research projects.

Therapist 4

TABLE 30 Summary of audio-recorded evidence of delivery of core components by group and by appointment

Appointment	Treatment group			
	Basic PFMT	Biofeedback PFMT		
	Basic core components	Basic core components	Biofeedback core components	Basic and biofeedback core components
Appointment 1				
Potential range	0–17	0–17	0–7	0–24
<i>n</i> (audios)	7	8	8	8
Median	11	8.5	3	11.5
Mode	10	8	2 and 3	15
Range	9–14	5–12	0–7	7–18
Appointment 2				
Potential range	0–15	0–15	0–11	0–26
<i>n</i> (audios)	6	6	6	6
Median	7.5	6.5	4.5	10.5
Mode	10	7	4 and 5	9
Range	6–10	3–8	3–8	7–15
Appointment 3				
Potential range	0–16	0–16	0–9	0–25
<i>n</i> (audios)	8	5	5	5
Median	7	9	4	11
Mode	5, 7 and 8	6 and 9	0	Multiple
Range	3–12	6–11	0–7	6–16
Appointment 4				
Potential range	0–17	0–17	0–10	0–27
<i>n</i> (audios)	6	6	6	6
Median	6	6.6	5.5	11.5
Mode	5	4 and 7	6	10 and 13
Range	4–10	4–8	3–7	7–14
Appointment 5				
Potential range	0–14	0–14	0–10	0–24
<i>n</i> (audios)	8	5	5	5
Median	8.5	6	6	12
Mode	9	6	6	12
Range	4–11	1–8	1–6	2–14
Appointment 6				
Potential range	0–16	0–16	0–5	0–21
<i>n</i> (audios)	11	12	12	12
Median	7	7.5	3	11
Mode	7	8	2 and 3	10.5
Range	5–9	2–10	1–5	5–13

Some centres retained trial monies to use for staff training and development, whereas sometimes the hospital or trust kept the money; however, all centres kept their laptop and biofeedback units at the end of the trial. One centre's IT department required them to buy their own laptop at £1000 (the OPAL trial contributed the value of the trial laptop) and the therapist from this centre noted that the subsequent per-patient-recruited payment helped them cover OPAL trial set-up costs. The remaining therapists were not motivated by the payment but by the research, a desire to know the outcome and acquiring biofeedback units (or extra units) in the department, or the staff training and development:

I mean, obviously, what I'm hoping is that this is going to show that biofeedback is better than no biofeedback, because so many of my patients find it, they just love it, they absolutely love it, it just changes the way they see their appointments, and it also, I, I just feel it empowers patients to do things for themselves.

Therapist 13

... we don't actually have any [biofeedback units] in the department; we tend to suggest that patients go and purchase their own, we don't have a pool of them, so that'll change now.

Therapist 5

Across all centres, those involved in recruitment reported that some participants refused because they could not commit to six appointments. Others were reported to join for altruistic reasons or because they wanted the 'new' therapy, even though the therapist made it clear they had only a 50% chance of being randomised to the biofeedback PFMT group.

Delivering the trial

Therapists were positive about the intervention delivery training. A main concern was insufficient hands-on experience with the biofeedback units, including connecting it to the laptop; this was true even for therapists with experience of biofeedback devices from other manufacturers. It was recommended that therapists practised with the units after the training and, although some did do this, many of the interviewed therapists would have been willing to try the units on each other as part of their training. Another concern was the delay between training and starting trial intervention delivery. Therapists would have liked more timely training or a refresher nearer the start of the trial on biofeedback, IT and paperwork, and such refreshers were offered and received by many. The trial website was reported to be easy to use and navigate.

All centres experienced delays (a few months to 1 year) in being able to commence intervention delivery, as the OPAL trial-supplied laptop was taken away by their IT departments for NHS security software installation (one centre had to procure an alternative laptop). Therapists spent considerable time chasing IT departments to get trial laptops back. Once ready to start, a variety of further IT set-up problems were experienced by most centres, at least initially, while they got familiar with the equipment. Problems included the software installed by NHS IT and faulty equipment: biofeedback units, vaginal probes, fibre optics and dongles. Often the therapists who experienced these issues had to address these IT problems by themselves. They reported that a lot of time was spent going between the trial team, their IT department and the manufacturers to get to the bottom of their problem. Although the trial team support by telephone was perceived as good, for those working on their own, this was a frustrating and lonely experience, and these IT problems were the leading reason why one centre withdrew from the trial.

All therapists found the trial paperwork lengthy, time-consuming and repetitive, particularly the checklists. Some said they could not complete the paperwork within the allocated time and some tried to complete everything and realised later some elements (e.g. optional components) were not necessarily meant to be repeated at every appointment. Some therapists felt that the paperwork did not ask for a subjective history, which they would normally cover in their routine clinical work, especially before vaginal examination.

However, there was free-text space for them to add anything. 'Protocolised' TAFs and checklists were sometimes counterintuitive to the holistic, tailored approach therapists typically used:

I try and have quite a holistic approach to things, and I don't always feel like the pelvic floor exercises are the first thing that I would tackle, so sometimes I might, you know, talk to people about you know, more like fluid management stuff and it would be that I wouldn't even get started on the pelvic floor stuff until I'd sorted out whether or not they have an overactive bladder for example.

Therapist 3

Although two centres bought extra therapist time to cover additional OPAL trial work, most centres had to absorb this into existing workloads, even though they were not being asked to see additional patients. Three centres withdrew from the trial, as the therapists did not have capacity to deliver the trial or they found the demands of trial delivery overwhelming and lost motivation to take part. A combination of factors contributed: difficulties recruiting women, IT or biofeedback device problems, paperwork issues, changes or loss of key staff and delays between centre training and the centre being ready to deliver the trial. Three further centres stayed in the trial but struggled to recruit because of caseload mix (few eligible referrals and complex clinical presentations taking trial-designated therapists away from trial-eligible patients), lack of staff continuity (e.g. rotational junior staff, maternity leave, staff leaving) or IT issues (which delayed start of recruitment).

Delivering the basic PFMT and biofeedback PFMT interventions

All therapists liked the basic PFMT intervention, describing it as their 'bread and butter' (therapist 12) and 'what we would have been doing anyway' (therapist 1), although several said that the number of appointments was more than they would routinely offer:

... we probably saw them more often, because we were seeing them the six times, so, on average, I'd probably see a patient probably two to three times, so they probably were getting more contact with us, which I think, from a motivational point of view, probably is better.

Therapist 27

The basic PFMT intervention was delivered within the allocated time and some said they struggled to fill the six appointments or participants were not sure why they had to keep attending. Some therapists thought components related to dietary advice and fluid intake were more often delivered in the basic PFMT intervention.

All therapists also liked the biofeedback PFMT intervention, but it was more problematic to fit it into the scheduled time because there were additional activities (e.g. downloading biofeedback data, showing participants how they had improved and adjusting the unit). Many therapists experienced problems initially in the delivery of biofeedback, particularly connecting the biofeedback unit with the laptop. Already feeling under time pressure in a biofeedback consultation, this caused additional stress and these early trial consultations ran late:

Very stressful; my heart used to sink a little bit if they were randomised to the EMG group, I just knew that that was going to be a lot more work and a lot more paperwork and it probably would be appointments that would run over time ...

Therapist 12

Although many technical problems resolved over time, having prior experience of using biofeedback was useful. Some centres routinely used similar hand-held biofeedback units, whereas others had only one hand-held unit for the whole department or a larger clinic-only machine. Previous and/or regular experience of using hand-held units gave therapist's confidence in delivering the OPAL trial biofeedback intervention, which they continued to use post trial. In contrast, confidence was usually lower in those without previous experience and, although encouraged to practice with their routine patients (prior to the trial), the less experienced still wanted more time to practice before delivering biofeedback to trial participants.

Although the basic PFMT intervention was similar to usual practice, the biofeedback protocol was different for many:

... it very much differed from my usual care ... in as far as I wouldn't, I have, I haven't given any equipment to aid the patient's compliance before.

Therapist 26

The trial protocol required therapists to introduce biofeedback in clinic at appointments 1 or 2, then at every subsequent appointment, and prescribe home biofeedback. In contrast, therapists who previously used biofeedback said that they would typically use it more selectively, for instance they would not usually start a patient on biofeedback immediately but start with PFMT, or would introduce biofeedback only to patients who struggled to contract the pelvic floor muscles, or only once patients had good PFMT technique to check that they were not using accessory muscles (albeit with different units to those used in the OPAL trial).

Delivering treatment beyond the trial

Some therapists delivered treatment to participants beyond the 6-month blind assessment date, explaining that more improvement was possible. Two interviewed therapists extended treatment (after the six OPAL trial appointments) for participants in the basic PFMT group, to include biofeedback. Therapists also referred participants for further treatment or investigation according to their clinical judgement.

Therapists were more confident using biofeedback in their everyday practice post trial. They were also using it for a wider range of patients:

I probably used it a bit more often, so the OPAL randomisation maybe made me use it with people that I wouldn't usually have used it with.

Therapist 14

The randomised allocation required therapists to use biofeedback with some patients they might not usually have offered it to and their assumptions about who might find it acceptable had changed. For example, they had typically thought younger women would like the technology and older women would not, whereas it became apparent during the trial that younger participants (often working mothers) did not have the time or privacy to use the biofeedback at home, whereas older participants did and the technology in itself was not a barrier to them.

Therapist perceptions of participants' intervention fidelity and adherence

Therapists made a number of theoretical propositions about participants' engagement in treatment and adherence. One observation was that symptoms prompt participants to undertake PFMT and that participants forgot to do PFMT as their symptoms improved.

A second observation was that participants' levels of motivation and commitment to the intervention influenced engagement and, ultimately, outcome. For instance, the importance of patient motivation and buy-in was linked to participants' health priorities (many participants in the trial had comorbidities and their UI symptoms were not their top health priority) and whether or not participants could find energy and time for PFMT. Another example was that therapists thought that if participants expected the biofeedback to stimulate their muscles for them, they were disappointed and somewhat demotivated. Alternatively:

... if they need motivation or they're very motivated, and particularly if they're very young and they have less complex problems, if it's just the pelvic floor issue then you've got more time and energy to focus on, right, pelvic floor, biofeedback, whereas if it's, you know, you've got to be dealing with their constipation, their diet, their depression, or a million other things, you're less likely to give up a whole half-an-hour appointment just for the pelvic floor.

Therapist 8

Therapists also saw buy-in change over time, with participants withdrawing from the trial because of personal circumstances (such as ill health or ill health of a family member), rather than because of the research or the interventions. These participants prioritised their other health conditions or the health of their family member over their continence.

Accountability was also thought important. Therapists mentioned that participants compared treatment to weight loss classes; regular attendance and knowing they were going to be assessed meant that participants worked harder than they would have done had they been on their own. The therapists wondered, but were not sure, if accountability was felt more strongly in the biofeedback PFMT group, as part of the appointment included the therapist downloading the participant's data.

Therapists noticed how many participants struggled to fit biofeedback into their daily routine, especially those who were time-constrained and working mothers:

... how do I fit this into my daily life? – that's the big issue, and you know, we had quite a few conversations about that.

Therapist 21

Therapists wondered if biofeedback would work better if it was used only in the clinical setting rather than at home, or if home biofeedback was the key, or if biofeedback would be appropriate only for those women who struggle to do PFMT because of very weak pelvic floor muscles or a lack of contraction sensation. The majority of interviewed therapists hypothesised that biofeedback was more suited to goal-orientated women who have the time and privacy at home to use the equipment:

I did also probably feel that maybe more active sporty ladies who had quite high demand on their pelvic floor and quite a high demand from treatment wanted, you know they, they seemed to be better with the biofeedback, I definitely felt like it's a motivational tool and made ladies feel like they were achieving something, and made them work harder probably.

Therapist 30

Participant diaries

Although the number of returned diaries dropped from appointment 2 to 6, in much the same way appointment attendance dropped (see *Table 30*), there were no apparent differences in the pattern between the two groups (*Figure 11*). Nor were there apparent differences in the proportion that were signed by participant and therapist (a BCT called 'commitment').

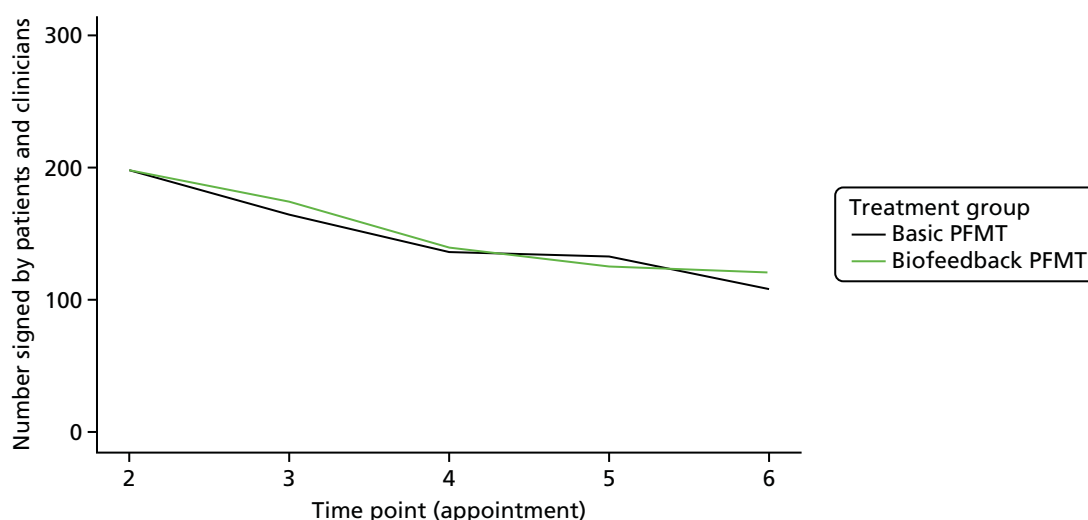


FIGURE 11 Proportion of exercise diaries signed (by participant and therapist) and returned by appointment and by group (in respect to the total number attending that appointment in that group).

In free-text diary entries from participants, the most frequent reasons for being unable to exercise were time (biofeedback PFMT = 13 comments, basic PFMT = 33 comments), forgetting (biofeedback PFMT = 7 comments, basic PFMT = 24 comments) and other physical health reasons (biofeedback PFMT = 25 comments, basic PFMT = 13 comments), with menstruation featuring more frequently in the biofeedback PFMT (14 comments) than the basic PFMT group (6 comments) (participants were not expected to use biofeedback during this time).

Discussion

Discussion is structured according to the source data, covering key findings and interpretation, as well as the strengths and limitations.

The appointment checklist data show that therapists said (ticked 'yes') they had delivered most of the core components per appointment; therefore, the between-group differences in delivery were as designed: the biofeedback PFMT group received more core components (specific to biofeedback) without necessarily compromising the number of basic PFMT core components the participants were meant to get. This suggests that, overall, the therapists delivered the interventions with fidelity to the form of the protocol for both groups and that the biofeedback PFMT group did receive an intensified intervention. It was clear that the optional components were seldom used. The strength of the checklist data set is that it may offer a comprehensive assessment of appointment content. The main limitation is that it relies on accurate therapist recording. Other data sources indicate that therapists were clearly time pressed (see *Therapist interviews*), so reporting accuracy may have been compromised and in the much smaller sample of audio-recorded appointments less was heard to be done. However, there was no suggestion that fidelity was lower in one group than the other.

The audio-recording data verified that the biofeedback PFMT group did receive more components (biofeedback specific) overall, and there was no obvious imbalance in number of basic PFMT core components between groups, suggesting, again, that the biofeedback PFMT group received an intensified intervention. We plan to compare individual audio-recorded appointments with their corresponding checklist to verify fidelity. This is a potential strength of the audio-recordings, yet they have limits as a gold-standard method of assessing fidelity; we cannot be sure that we coded every component of an appointment as some could have been completed in silence, outside the clinic room or after the recorder was turned off. Coding was complicated because therapists did not necessarily use the language or specifics detailed in the intervention manual or BCT taxonomy, and there was general conversation between the therapist and the participant (e.g. for establishing the therapeutic relationship) to 'sift' through. The resources required to obtain the audio-recordings, transcribe and analyse them was a further limitation. However, they may provide a rich data source of anonymised composite examples for future training to help therapists refine their use of BCTs.

The therapist interviews provided extensive material regarding the challenges of trial delivery and how this had an impact initially; they also offer further insights into intervention delivery fidelity and therapists' ideas about what has an impact on patient adherence. Several of the challenges expressed by therapists were to do with the research (IT issues, paperwork) and these would not be an issue if either of the trial interventions were implemented into routine practice. It was clear that some therapists were overwhelmed by the logistics of getting ready for the trial and there were gaps in support locally, highlighting how difficult it can be to establish research 'readiness', especially in centres where levels of research experience are low. Many therapist comments related to their confidence in delivery of good basic PFMT in contrast with their concerns about confident and competent use of biofeedback PFMT, at least initially (e.g. the 'heart sink' some experienced if a participant was randomised to biofeedback). Much of this concern related to IT issues and the time pressure imposed by trying to fit all biofeedback components in the first two appointments, rather than a dislike or lack of belief in the use of biofeedback PFMT or, indeed, the biofeedback protocol. Nevertheless, the therapists did manage to deliver an intensified intervention to the biofeedback PFMT group.

In contrast, therapists reported that both they and the participants felt that the basic PFMT intervention lacked content, with some participants querying why they needed to attend all six appointments. Checklist data show that few optional intervention components were used in either group. This is important, suggesting that the trial was a fair test of PFMT intensified by biofeedback and we can be clear that the null result is unlikely to be confounded by more use of 'extra' therapy (the optional components) in the basic PFMT group. However, from a clinical perspective, it is perhaps a missed opportunity when there were further optional components that could have been used to promote participants' adherence, or to overcome barriers to adherence. This raises the question of how well the therapists understood and adopted the theory underpinning the intervention protocol; they clearly distinguished between the basic and the biofeedback PFMT interventions, but it was less apparent that they had assimilated into their practice all of the underlying theoretical components (IMB and BCTs) for enhancing participant adherence to treatment. This suggests an opportunity for further refinement to the training manual (and more time for training) for promoting PFMT adherence. We do, however, commend the trial therapists for delivering the core components of a highly protocolised evidence-based basic PFMT intervention with such commitment. It may be that this excellent basic PFMT meant that the biofeedback had to add a lot extra to obtain additional effect on outcomes.

Therapists hypothesised that many of the influences on participants' PFMT (non-) adherence related to the complex context of participants' lives, similar to findings of Hyland *et al.*⁵⁷ These theoretical propositions can be explored in further analyses in which we link main trial, case study and process evaluation data sets and attempt to identify any key characteristics about the participants, their therapist or the intervention they received that influence outcome (see *Chapter 7*).

The exercise diary data indicate that one of the BCTs (commitment, indicated by participant and therapist signature) designed for promoting adherence was utilised consistently for both groups. Therapist interviews suggested that this was seen as a useful addition to the therapist 'tool box' and therapists attending the post-trial collaborators' event described continuing to use this BCT in their current practice (and that they had found the protocolised checklists useful reminders). The participants' free-text comments in the diaries confirmed some of the difficulties experienced when using vaginal probes; other physical health problems and menstruation appeared to affect biofeedback use more than on undertaking basic PFMT. However, time and forgetting were more frequently reported reasons for non-adherence in the basic PFMT group and a possible interpretation here is that having a biofeedback unit at home (whether used or not) meant that fewer participants forgot to exercise.

Conclusion

The process evaluation offers two key findings concerning fidelity of delivery and receipt. Together, they indicate that the OPAL trial did achieve what it set out to do: a fair test of whether or not a PFMT intervention intensified by biofeedback could improve participant outcomes. First, therapists did deliver an intervention to the biofeedback PFMT group that was more intensive than that delivered to the basic PFMT group; this was despite therapists being time pressed to complete the biofeedback intervention delivery and all the trial requirements. Second, most participants did receive core components of basic PFMT and this was the case for both groups. Few optional intervention components were used by therapists in either group and there was no inadvertent 'intensification' of the basic PFMT group intervention through an imbalance in the amount of optional components used, even though therapists reported that they sometimes struggled to fill appointments in the basic PFMT arm.

This embedded and comprehensive theoretically informed mixed-methods process evaluation is unique to the trials conducted so far in the field of PFMT. To the best of our knowledge, only one trial¹¹ has reported a process evaluation and this was a single-method (qualitative) approach; one new trial is planning a mixed-method evaluation⁵⁸ and an earlier trial reported plans for a process evaluation.⁵⁹ The multiple data sources in the OPAL trial bring a richness to our understanding of trial processes and experiences of the therapists

and women involved, providing an opportunity to explore factors having an impact on adherence to PFMT, and other contextual factors that may explain variations in treatment delivery and effectiveness across and between the trial groups.

The work is not without limitations in terms of data quality, analytical volume and complexity; a programme of further more nuanced analytical work is being prepared to ensure that we fully explore the opportunities for learning from these data sets. In summary, we have demonstrated robust assessment of intervention fidelity assisting in interpretation of the trial outcomes (as the null result is a clinically important message) and indicated potential transferability of these findings into clinical practice and further research.

Chapter 6 Longitudinal qualitative case study

Introduction

This chapter reports the methods and findings from the longitudinal qualitative case study. In line with contemporary process evaluation guidance, this is an in-depth, pre-planned and theoretically driven longitudinal, comparative, qualitative case study to support understanding of two complex interventions that aim to reduce UI in women.⁵³ In this chapter, we refer to the interview participants as women, in recognition of the fact that this chapter is based on women's interview accounts (rather than using the term 'participants' as elsewhere in the report).

In this chapter, the longitudinal qualitative comparative case study will be referred to as the case study. Given the link between this study and the main OPAL trial, the same conventions in terms of referral to group allocation will be adhered to: specifically, when referring to the basic PFMT group we are referring to women allocated to basic PFMT group (ITT), whether or not the women adhered to treatment or crossed over treatment group; similarly, when referring to the biofeedback PFMT group, we are referring to women allocated to the biofeedback PFMT group.

When quotations are presented, they are followed by the case number of the woman, the interview (0M for baseline, 6M for 6 months, 12M for 12 months and 24M for 24 months) and the woman's group allocation.

This chapter addresses one aim from the OPAL trial, namely to:

- investigate women's experiences of the interventions, identify the barriers and facilitators that affect adherence in the short and long term, to explain the process through which they influence adherence and to identify whether or not these differ between randomised groups.

Methods

A longitudinal, qualitative, two-tailed case study design⁶⁰ was employed, in which the tails were the biofeedback PFMT and basic PFMT trial groups. A detailed protocol has been published.²¹ A sample of women from both groups took part in semistructured interviews. The two-tailed case study design complemented the trial design in its comparative focus, with the analysis set up to explore group differences. In this chapter, we will hereafter refer to groups rather than tails, in line with the terminology used in the trial. Case study design supports robust group comparison in a qualitative way;⁶¹ therefore, conclusions of similarity and difference should be read as qualitative comparison as opposed to quantitative (statistical) comparison.

Sampling and recruitment

Forty randomised women (20 in each group) were purposively sampled for variance in centre type, women's type of UI and therapist type. Each recruited woman was one case. Women were asked to consent to the case study specifically (having already consented to take part in the trial). The women were given an additional invitation letter [see the project web page: www.journalslibrary.nihr.ac.uk/programmes/hta/117103/#/ (accessed 29 July 2019)] and patient information leaflet [see the project web page: www.journalslibrary.nihr.ac.uk/programmes/hta/117103/#/ (accessed 29 July 2019)]. Women who remained interested were contacted by telephone approximately 1 week later to ask if they would like to participate. Written consent was obtained at the time of the first interview [see the project web page: www.journalslibrary.nihr.ac.uk/programmes/hta/117103/#/ (accessed 29 July 2019)].

Case study data collection

Data were collected by a series of semistructured interviews [see the project web page: www.journalslibrary.nihr.ac.uk/programmes/hta/117103/#/ (accessed 29 July 2019)]. Each interview had a specific focus:

- Baseline pre-treatment interviews (face to face) explored the woman's experience of UI, the social contexts within which she experienced UI and her expectations of treatment.
- A 6-month post-treatment interview (face to face) explored the woman's experience of the trial intervention, her adherence to therapy appointments and the prescribed programme, factors that affected that adherence and her perceptions of treatment outcome.
- 12- and 24-month interviews (telephone) explored, at each time point, the woman's experience of UI post intervention, the intervention, factors that influence ongoing PFMT adherence and treatment outcome.

Interview data were, with consent, collected using a password-protected audio digital recorder. Interview audio-recordings were anonymised, transcribed verbatim and entered into NVivo software to support analysis.

Case study data analysis

Analysis was guided by the OPAL trial protocol [see the project web page: www.journalslibrary.nihr.ac.uk/programmes/hta/117103/#/ (accessed 29 July 2019)] and the OPAL qualitative study and process evaluation analysis plan [see the project web page: www.journalslibrary.nihr.ac.uk/programmes/hta/117103/#/ (accessed 29 July 2019)]. Three different researchers have worked on the OPAL case study (Anne Taylor, Aileen Grant and Marija Kovandzic), alongside the responsible grant holders (Carol Bugge, Jean Hay-Smith and Sarah Dean). By the nature of qualitative analysis, each analyst had a different approach to data analysis. This was encouraged by the grant holders, within the confines of the protocol, to maximise the insights into the data. Sources that were drawn on to support that analysis included Yin,^{60,61} Alvesson and Skoldberg,⁶² Grant *et al.*,⁶³ Kovandžić *et al.*,⁶⁴ Stake⁶⁵ and Ritchie *et al.*⁶⁶

Overall, analysis was iterative with data collection. Analysis occurred on four interacting levels to facilitate within- and cross-case comparisons.

At the level of the individual interview

An initial *a priori* coding scheme was developed and initially applied, focusing on core areas of interest: UI experience, PFMT ± biofeedback experience, factors that influenced adherence in the short and long term and perceptions of treatment outcome. The coding was developed through team discussions, iterative coding and multiple analysts' perceptions. The analytic purpose was to identify barriers and facilitators that influenced adherence and patient-reported UI outcomes.

At the level of the case (woman)

Case summaries in narrative and tabular form were written with a focus on understanding a woman's experience of UI, the treatment, adherence, treatment outcome and how these factors interacted. Analysis focused on identifying issues relating to changes over time and in developing rival explanations (additional theoretical propositions) that guided subsequent analysis.⁶⁰ Theoretical propositions and rival explanations are analytic strategies drawn from case study design.⁶¹ The theoretical propositions used in the OPAL trial were drawn from the original research questions and the rival explanations arose from working with the data.

At the level of the trial group

Using case summaries and matrices from the framework approach,⁶⁶ the cases for one trial group were arranged together and consistencies and inconsistencies searched for. The aim of analysis was to identify the core barriers and facilitators within the trial group, the detailed explanations for them and interactions between them.

At the group comparison level

The biofeedback PFMT and basic PFMT groups were compared using the theoretical propositions in order to identify similarities and differences in barriers and facilitators between the trial groups.

After the trial result was known, an additional analysis was undertaken that aimed to explore who biofeedback works for and why. This analysis is not presented in this report, but may be helpful in understanding subgroups of women for whom biofeedback is more useful.

Management and governance

Ethics approval for the case study was gained within the main trial approvals (see *Chapter 2*).

The case study and process evaluation team had a management group with the required mix of clinical, qualitative, quantitative and theoretical skills and experience. The group met regularly to discuss the research management and emerging findings. The case study was carried out at a separate academic institution to the main trial. The case study team participated in trial meetings to understand how the trial was progressing, but the case study and process evaluation team meetings were closed. Data were not shared from the case study and process evaluation group with the main trial group until the final PMG meeting in September 2018.

Findings

Sample

Forty women, 20 per group, were recruited to the case study, as planned. Twenty-five women completed all four interviews, but, owing to the technical problems with the audio-recorder, a full data set was available for only 24 women (10 biofeedback PFMT and 14 basic PFMT). The total data set consisted of 125 interviews, including 24 complete cases (96 interviews). The total number of minutes of recorded interviews per case ranged from 15 minutes to 126 minutes, with a total of 2856 minutes of recorded interview data. There were 40 baseline interviews (20 biofeedback PFMT and 20 basic PFMT), 32 interviews at 6 months (16 biofeedback PFMT and 16 basic PFMT), 28 interviews at 12 months (13 biofeedback PFMT and 15 basic PFMT) and 25 interviews at 24 months (11 biofeedback PFMT and 14 basic PFMT).

The age of women in the case study ranged from 20 to 76 years, with both the biofeedback PFMT and the basic PFMT groups including women with a wide age range (*Table 31*). In the main trial, women ranged in age from 20 to 83 years (22–83 years in the biofeedback PFMT group and 20–78 years in the

TABLE 31 Characteristics of women in the case study by group allocation

Variable	Treatment group		Overall case study sample
	Basic PFMT	Biofeedback PFMT	
Age (years), range	20–76	25–69	20–76
UI type, <i>n</i>			
SUI	6	5	11
MUI	14	15	29
Centre type, <i>n</i>			
Community	2	4	6
University	9	7	16
DGH	9	9	18
Therapist type, <i>n</i>			
Physiotherapist	18	18	36
Nurse	2	2	4

DGH, district general hospital.

basic PFMT group); thus, the women in the case study were comparable in age to the main trial sample. From the total case study sample, 11 women had SUI and 29 MUI; the proportions were similar within groups. Six women in the sample were treated in community clinics, 16 in university hospitals and 18 in district general hospitals; again, there were similar proportions in the groups. The vast majority of women were treated by physiotherapists ($n = 36$) and four women were treated by nurses.

Women's adherence to the interventions

Women's adherence to the interventions was analysed in two phases: active treatment and maintenance. 'Active treatment' refers to the time when women were attending appointments and receiving the OPAL trial interventions delivered by a trained therapist. It is the proxy for shorter-term adherence – the uptake and adoption phase of PFMT – including women's attendance at appointments, receiving biofeedback-mediated PFMT or basic PFMT in the clinic and then undertaking their prescribed programme (biofeedback PFMT or basic PFMT) at home between appointments. The 'maintenance' phase is when long-term adherence is demonstrated and is the period of time after the active treatment has ended when women were asked to continue PFMT themselves at home without therapist supervision, including relapse management, up to their final follow-up at 24 months.

Table 32 shows examples of the variation in women's adherence to PFMT. These examples illustrate that there are no obvious group differences in adherence in the case study sample in terms of the frequency with which they undertook biofeedback PFMT or basic PFMT.

Facilitators of adherence during the active treatment phase

There was greater similarity than difference in facilitators of adherence in the active treatment phase when the trial groups were compared. Two key themes, among the many that were identified, focused on UI symptoms and factors related to the OPAL trial therapist.

Urinary incontinence symptoms acted as a facilitator in several ways. One way was through the mechanism of women wanting to eliminate or reduce their UI, so that they could get on with their lives and improve their quality of life:

Case 27, 0M, biofeedback: Well I'm hoping that it'll help the leaking and it'll, it might never stop, but it won't be as bad as it's been . . . that's what I'm hoping.

Researcher: Yeah. Is there a goal; do you have, like, a personal goal that you would like?

Case 27, 0M, biofeedback: Just that really, just . . . to stop the leaking, maybe be able to go back to yoga and not feel like I'm worrying about leaking or whatever.

TABLE 32 Case study examples of variation in adherence to treatment by allocated treatment group and across time

Adherence	Active treatment		Maintenance	
	Biofeedback PFMT	Basic PFMT	Biofeedback PFMT	Basic PFMT
Good adherence	Case 27: uses biofeedback every couple of days and has exercised consistently with no breaks	Case 14: undertakes PFMT 'religiously'	Case 17: does PFMT at least daily	Case 38: does PFMT at least daily
Moderate adherence	Case 39: very good adherence for first couple of months then more ad hoc	Case 24: does PFMT 'most of the time'	Case 1: tried for three times a day but does not always manage – does short pulses and not long holds	Case 36: does PFMT when symptoms return
No/minimal adherence	Case 2: maybe exercises one a week	Case 19: does PFMT irregularly	Case 32: does not do PFMT at all	Case 15: does not do PFMT at all

I'm not that old that I, I'm ready to kind of hang up my dancing shoes.

Case 26, 0M, basic

Women also wanted to prevent a deterioration in their UI symptoms and to avoid surgery. Seeing an improvement in UI during the active treatment phase motivated women to adhere because they felt that their treatment, and their skill to undertake the exercise, was working:

Doing the exercises [was most helpful about treatment] and noticing that there was a change, do you know what I mean? And then realising myself that that was, there had been a change . . .

Case 24, 6M, basic

For women from both groups who had a break in their regular biofeedback PFMT or basic PFMT practice, a deterioration in symptoms (after a period of improvement) provided proof of PFMT effectiveness and acted as a facilitator to use the skills that they had learned to overcome the symptomatic deterioration.

Many women from both groups talked at length about the positive impact of the therapist. Women talked about their therapist as an important and credible source of information, as a motivator and as someone who taught them the exercise, lifestyle and behavioural skills needed to undertake biofeedback PFMT or basic PFMT (in line with theoretical model underlying the interventions).¹⁷ All of these factors influenced adherence in the active treatment phase in both groups. However, possibly the most important element of the interventions in each trial group was the instruction on how to perform PFMEs, given by the therapist during the vaginal examination (digital assessment). Given the sensitivity of the topic, vaginal examination was not easy to talk about during the research interviews and, consequently, not an easy finding to capture in the analysis. Yet there was a consistent observation of the importance of the therapist-mediated vaginal feedback as being one of two distinctive and valuable forms of vaginal feedback in PFMT (therapist mediated and EMG mediated). The findings from the case study point to the therapist-mediated feedback as being the priority and as one of the most important therapist-related facilitators in gaining confidence in PFMT skills and adhering to treatment.

The quotation below provides an illustration of the difficulties of articulating experience of PFME instructions during the vaginal examination, as well as the importance of these instructions, which included feedback (as exemplified in the quotation, a part of the feedback loop was the act of the therapist feeling the difference in muscular activity during the examination):

That was quite good actually, having somebody there, and I think when you're doing exercises and then being able to feel that it was working, do you know that way when you would get your assessment . . . and you did have to do them, the exercises, and she could feel the, the difference [. . .] I felt [. . .] that was good, u-uuh, just to know that you were doing it properly [. . .] 'cause you do those exercises and you really don't know one way or another if you are doing it right.

Case 30, 6M, biofeedback

Another important therapist-related factor that had an impact on adherence was the rapport created between the therapist and the woman. The conditions for creating rapport require further analysis. It is possible that the above-mentioned therapist-mediated vaginal feedback plays a role, but at this point of analysis it is certain that having dedicated space and time (secured by OPAL trial intervention design) to build understanding and trust through repeated appointments with the same therapist acted as a motivator to adhere to the treatment, if not being a therapeutic agent on its own:

And it's very motivating . . . you know, seeing someone who's interested in you, who wants to help you is terribly motivating . . . 'cause otherwise you're just on your own, 'cause you don't chat to your friends about it . . . the only person I've ever really spoken to about this [UI] is [OPAL therapist] and the nurse specialist.

Case 32, 6M, biofeedback

[. . .] it was good having a one to one with someone who kind of constantly, you were able to talk to about your symptoms and how to improve it and I think just knowing that em that they were there and they were able to tell you, you know, 'if you work on this, it will improve' and I think that was a big help, even right at the end there, it was a really good for her to tell me, the physio[therapist] to tell me, like, what exercises it's best for you to do, what's not good for you to do and if you keep going wi' this it's going to continue to improve, I think that will help me. [. . .] you know, psychologically, even if it wasn't physically, you know, I mean it eventually will be physically, but em, you know, even psychologically I think that was good.

Case 15, 6M, basic

Other therapist-related facilitators included education provided by the therapist, being treated by an accommodating and skilful therapist, being treated by a therapist who adjusted the treatment protocol based on individual needs and feeling accountable to the therapist:

I think it was the, the, eh what's the, what's the best way to describe it, the actual having to report back to [therapist], because then you knew, you know, you can't, well you can't just sort of, you know, sit there and say 'right, OK I didn't do it,' and she would know herself when we did the sort of, the few, even, you know, not the internal examination, when we did the actual work, you know, when she was there and she could tell from my posture, you know, if I was doing it right or not, she was like 'right, you're slacking' . . .

Case 3, 24M, basic

Beyond the symptom- and therapist-related facilitators summarised above, women identified other facilitators of adherence that included the following:

- Service structure, framing and physical environment. Having regular appointments; ease and flexibility of making appointments; feeling positive about the physical environment of the treatment facilities; feeling that the intervention was within the framework of womanhood; or the woman finding the treatment as a whole a novelty were all facilitators of adherence.
- Ownership and agency in the intervention. In terms of ownership and agency, the following were seen as facilitators of adherence: the woman feeling accountable to herself; having a new focus on herself through starting a new lifestyle routine; developing a sense of embodiment and empowerment as starting to feel (and control) her own pelvic floor muscles; and through their own personal characteristics, such as 'strong will', commitment and altruism:

I was so determined though, I mean the thing is you've got to want to, to help yourself I think [. . .] you know, it's just not going to, just taking a note of what somebody says to do, you've got to want to do it as well [. . .] you've got to want to, you've got to need to do it as well, you know.

Case 20, 24M, basic

- Support from relevant others. Their partner, participation in the trial and a sense of accountability to the trial team were all facilitators of adherence.
- The intervention. In terms of the intervention, the following were seen as facilitators of adherence: making use of the exercise diary; a woman gaining knowledge about her anatomy, UI and PFMT; learning the skills to undertake biofeedback PFMT or basic PFMT; learning routines for biofeedback PFMT or basic PFMT (such as keeping the biofeedback unit next to her bed) or doing PFMT while watching TV; and learning that PFMT exercise was easy and could be performed at any place and any time:

. . . so the education was eh the principal thing, when you learn how to do and why it's wrong, what is wrong . . . and then you can do your, do good for your body.

Case 13, 6M, biofeedback

... it's probably the easiest form o' exercise you could do, I mean you don't even need tae go tae a gym, it's so easy.

Case 34, 12M, basic

There were facilitators that were specific to biofeedback. Some women reported liking the biofeedback device and having confidence that, by using biofeedback PFMT, they were more likely to achieve symptomatic improvement than if they were doing PFMT alone. In developing the OPAL trial intervention, the research team hypothesised that visualisation of the pelvic floor muscle contraction via biofeedback PFMT would support self-efficacy for performing the correct contraction, leading to improved adherence and better outcomes. Some women in the biofeedback PFMT group did report that visualisation was important for them for two main reasons: (1) they could see if they were doing the pelvic floor muscle contraction correctly and (2) they could see improvement in their pelvic floor muscle contraction ability over time. Women valued the opportunity to be able to discuss the visualised contraction with their therapist.

Other features of biofeedback PFMT that women valued were biofeedback supporting women being competitive with themselves; having a new 'toy' to play with; the physical presence of the unit acting as a reminder; getting instruction from the biofeedback device in terms of counting of repetitions and pace of PFMT; and an awareness that the data on the biofeedback device would be looked at by and discussed with the therapist:

I thought it was quite positive that when you were actually using it you could see, and I think it did make you try, it definitely made me try harder, and also I felt that I was doing it for longer, like it, you know, a 10-second hold I think when you haven't got the biofeedback is probably, in reality, an 8-second hold, because you count quite quickly ... whereas with the biofeedback I felt that you were doing it properly and I was definitely trying harder because I was seeing it and I was thinking 'right, I want' it's that sort of slightly competitive side to human nature, you're thinking 'right, I want to get, I want to get it higher'.

Case 8, 6M, biofeedback

In summary, although some group differences were noted, there were more similarities in facilitators of adherence than differences. There were many facilitators of adherence in the active treatment phase, with being motivated to improve symptoms and the effect of the therapist being clear facilitators in both groups.

Barriers to intervention adherence during active treatment phase

There were more similarities across barriers than there were differences between the groups. Time and contextual factors in a woman's life (such as daily routines) were two of the themes that could be seen to act as barriers to adherence.

Women talked about having a lack of time for themselves; hence, finding time for appointments and to exercise was difficult. Women reported a lack of time to attend appointments in general and frequent appointments in particular, to focus on practising PFMT, either with or without biofeedback; biofeedback was even more time-demanding and, as such, a potentially greater barrier in the biofeedback group:

I don't know who supplied the physio[therapist] with the dates, but she kinda had a calendar, at the end o' my appointment she could tell me the time frame when I was due back ... and I would look at my diary and was like 'oh, that's only like 2 weeks' time', so I don't know, maybe even once every 5, 6 weeks or something, em but that, again, that's just because I'm a working mum and I don't always have the child care, so em it wasn't always easy for me to, to get the kids watched, ...

Case 16, 6M, biofeedback

Lack of time was compounded by having a generally busy life that included being a working mother, having unpredictable work patterns and going on holiday. For several women, their UI, and its treatment, was not a priority given the array of other things that were competing for their time. Illness – theirs, or in family members – was a particular barrier to adherence:

Em . . . most of the time I'm OK now, as I say I still do my pelvic floor exercise at the moment, eh it's not always OK, but [most of them are?], em [sighs] that's nothing to do wi' the machine [?] that I dropped out [of treatment], I took, mum took no' well and I took really bad depression and I would'nae get out the bed.

Case 17, 6M, biofeedback

Other contextual factors that acted to diminish adherence included not having a routine (or hook) for doing PFMT, lack of privacy at home, lack of support from their partner and simply forgetting (in the array of other things to do).

Several other barriers could also be identified, these included the following:

- A lack of sufficient, or sufficiently quick, improvement in the UI symptoms. This led to a drop in motivation to adhere. Despite this drop in motivation, many women were still inclined to continue treatment.
- There were service factors that acted as a barrier to adherence, some of which were OPAL trial-specific and others which were not. These included general transport or parking problems; the treatment facilities being unappealing or inappropriate; a lack of suitable appointments; and, although uncommon, a feeling that staff were being difficult, 'one of the receptionists at the [clinic] is an absolutely [sic] dragon' (Case 14, 6M, basic). Women also reported anxiety and/or embarrassment about having to undergo vaginal examination; being put off by other trial requirements, such as completion of the bladder diaries; not being randomised to their preferred group; and, in a few cases, there were indications of rapport and/or satisfaction with the therapist not being as good as for the majority of other women. Some women commented that there was repetition between the visits and that this made them unsure about the value of continued appointments:

Yeah, there seemed to be quite a lot, you know, I seemed to have a lot of appointments, em . . . my husband's going 'oh you're not going there again, what are you going for, what on earth are you going for this time?' . . . Em maybe the odd time I did feel a bit like that 'cause I felt, u-uuh, at times I thought 'oh God here, we're just going to talk about exercises' [. . .] the odd time I did feel 'gosh, maybe that was a bit of a waste of time' [slight laugh] . . .

Case 15, 6M, basic

- A few women in both groups felt that they did not learn the correct PFMT technique in the clinic or how and when it should be applied, which could manifest as a lack of confidence about doing the exercises correctly. It is possible that this lack of confidence is linked to not receiving enough therapist-mediated vaginal feedback:

I thought there would be, I thought there would maybe be more em involved in helping support you doing the actual . . . exercises; not that they were difficult or anything like that, I just, I, I think I just felt, you know, you get told what to do, you're advised about what, how to do them, they don't, [sighs] I've only once been checked to make sure I was doing them right, em so my feeling kinda was am I doing these right? Are they really effective?, and it was a bit hit and miss I felt . . . to how well I was doing; . . .

Case 26, 6M, basic

There were some barriers that were specific to the biofeedback PFMT group. Some women found the biofeedback device intrusive or painful to use and others found it inconvenient (e.g. having to set it up, or to clean it):

I found it intrusive and painful to be honest [. . .] if I had of [sic] found it less uncomfortable it possibly would have made me notice what I was doing more, but I, I just couldn't put up [with] the, the pain of it, so I couldn't be bothered with it.

Case 5, 24M, biofeedback

Women reported that they needed to find even more time to undertake PFMT supported by biofeedback. Some women also reported embarrassment and a lack of privacy about using biofeedback:

I think it was quite a good idea, but I don't think it worked for me, for my personal circumstances, I found it too footery [fiddly] to do, and I just found it quite difficult to have that kind of privacy . . . just to do it, because I found it easier if I was lying down in the bedroom but then, you know, the kids were always like in and out, running around and obviously I didn't want them to see it, and I just felt it took quite a lot of time and I just felt I didn't really have the privacy to do it properly, em, so I don't think it really worked for me, I felt it was too footery; but on the other hand I think it had lots of advantages, 'cause I think it was quite useful to see, to see what was actually happening.

Case 8, 24M, biofeedback

Other issues with the biofeedback included one woman reporting that she got thrush from using the biofeedback unit; the biofeedback unit could be framed as externalising the movement of the pelvic floor muscles and a distraction to embodiment of PFMT; and practical problems with the biofeedback unit that hampered ability to use it:

I thought I was doing super, then one day it died and it, I knew it had a brand new battery so that shouldn't have happened . . . it died, so I rang them up and I took it in and we got a new battery, then I came back and it happened again, it kept doing weird things, and then I bought batteries up the road in the end, so, . . . And then I realised that by looking at the machine I was distracted from doing the exercises.

Case 32, 6M, biofeedback

In summary, there were more similarities than differences in barriers to adherence in the active treatment phase; there were also additional barriers in the biofeedback PFMT group. A lack of time and many contextual factors were the key barriers to adherence to biofeedback PFMT and basic PFMT.

Facilitators of women's adherence in the maintenance phase

None of the women in the biofeedback PFMT group reported using biofeedback after the end of treatment in the trial. None of the interviewed women reported buying biofeedback equipment; some therapists did give women the probe to keep and use, yet none of the women reported using it, even though some reported intention to use it. Thus, the data below relate to women, from both groups, undertaking basic PFMT in the maintenance phase.

Women in both groups reported a change in their adherence from the active treatment phase. PFMT maintenance was not consistent over time in either group and there were no differences (from qualitative comparison) between the groups in their adherence. The inconsistency in adherence between women can be seen in *Table 32*, in which, at the extremes, some women undertook PFMT in a regular and daily manner, whereas others did not do PFMT at all. In between these extremes were women who undertook PFMT with varying degrees of regularity. As well as the inconsistency between different women, there were fluctuations in adherence for individual women over the time period with, for example, other health concerns taking over and diminishing adherence at some points in time.

Many of the facilitators that applied when women were in the active phase of treatment also applied in the maintenance phase.

Similar to the active treatment phase, women's desire to lessen UI symptoms supported adherence to PFMT in the maintenance phase. If women perceived symptom deterioration or recurrence and associated this with PFMT as a mechanism to improve symptoms, adherence was facilitated. The interpretation of the data would suggest that symptoms may only act as a prompt to undertake PFMT in the maintenance phase if the woman perceived that there was an improvement in symptoms as a consequence of doing PFMT during the active treatment:

Not really no [been doing PFMT], but quite often in the last week, 'cause I've noticed a difference that's why I've sort of started to try and do it again, 'cause I have noticed a difference in not doing it . . .

Case 8, 6M, biofeedback

Oh yes, I always will [exercise] now, that's it . . . that's it, because I know it, I know it's, you know how much it's helped.

Case 20, 24M, basic

There were multiple factors that seemed to influence women's confidence (self-efficacy) to continue, or feel able to restart, PFMT in the maintenance phase. Many women reported feeling that they had good levels of knowledge and skill to undertake PFMT correctly. Beliefs in their skills and knowledge could be attributed to women feeling they had mastery of PFMT; having memories of the support they received from the therapist during active treatment; recalling information imparted by the therapist; using the resources given by the therapist (such as information leaflets); keeping a record of PFMT like an exercise diary; and recalling the sense of hope given during treatment and the control they gained:

I don't feel like I need to go back and see a doctor or, you know, see a nurse or anything, I feel like if it got bad again I could, you know, I've got these exercises to fall back on.

Case 27, 24M, biofeedback

I remember the girl who, or the nurse that, the lady, you know the . . . pelvic floor . . . in [location], and I remember her, she was very good, gave me a lot of confidence in myself, you know and . . . it was really good, she was very, very helpful, and I can remember, I can remember the improvement, . . .

Case 20, 24M, basic

In the biofeedback PFMT group some women related having good skills and knowledge of PFMT directly to biofeedback during active treatment. Women in the basic group also felt that they had good skills and knowledge of PFMT acquired from teaching by, and feedback from, therapists. Therefore, biofeedback was not a necessary prerequisite for having skills and knowledge for PFMT maintenance:

[I remember] learning to use the machine properly . . . knowing I was doing it right and . . . yeah, and just generally being made more aware of the muscles that you need to squeeze and . . . when you're, and you know you do one at a time and then you hold them all . . . so yeah . . . being taught how to do pelvic floor . . . muscle training . . . yeah, being taught that properly, yeah, . . . made a big difference.

Case 23, 24M, biofeedback

Other factors that facilitated adherence in the maintenance phase included the following:

- A supportive home environment.
- Establishing the intervention as part of life: being able to find time for themselves; helpful work patterns (e.g. working from home or time spent commuting); making use of available time to do exercises (e.g. sitting and waiting time such as while commuting, or watching television).

- Development of routines or ‘hooks’ that continued to supported continued PFMT. In the example below the woman refers to keeping note of her exercising (like in an exercise diary):

I don't think so, I mean I know what to do, I mean it's . . . I, I, aye, I know what tae dae and I know what I should be doing but it's just the getting intae it, so . . . maybe I should start up a wee book kinda thing again, I done that the last time and I was dain' wee [unclear word] like when, how many I had done, do you know, how many kinda exercises I'd done that day, and eh, do you know what I mean, I think I should start that, when it's doon in black and white sometimes that kinda, kinda motivates you oan . . .

Case 24, 12M, basic

- Ownership and agency in the intervention – motivation and determination; cognitive features, such as remembering to do PFMT; and framing PFMT as ‘control’ (e.g. PFMT as a tool to control her body, or PFMT and the control over UI it gives being the only part of her life she can regain control of):

Oh probably [doing PFMT] daily, 'cause I do sort o' try to keep it going . . . 'cause I've got to keep control o' something, I can'nae control everything else [that I've got?] . . . [I was more?] conscious of it then, but, as I said, it's one thing I'm sort of trying to keep control of . . . so I'll try and keep that bit going.

Case 17, 24M, biofeedback

- Trial-specific factors – research interviews acting as a trigger to undertake PFMT, interviews providing a space for reflection on a woman's own PFMT practice, attending the 6-month pelvic floor assessment and wanting to demonstrate that the therapist had done an excellent job. These factors occurred in both groups.
- Women in both groups reported using an application on their phone that supported doing/remembering to do PFMT:

I [got] the squeeze App on my phone and that was really good . . . you know, it helps you, you obviously train yourself to hold for longer [kind of thing], that was good.

Case 19, 24M, basic

In summary, adherence did change in the maintenance phase from the active treatment phase. There was considerable variance, in individual women and between women, in adherence in the longer term. Many of the facilitators that supported women in adhering in the active treatment phase continued to facilitate adherence in the maintenance phase.

Barriers to women's adherence in the maintenance phase

One barrier to adherence that was unique to the maintenance phase was the loss of therapist support, and accountability to the therapist, when the active treatment phase ended. Some women felt ‘alone’ in their efforts to improve their UI. Others expressed the view that, because they were no longer accountable to the therapist, there was no longer that prompt to exercise. Other women said that they got out of the habit of writing their exercises down (as they would have done in the exercise diary during active treatment):

I thought, you know when the nurse did it with me, you know, did it, that helped me a lot, really it did, if I could keep going to the physiotherapist and if she kept checking me, because I think, you think you're doing it right and then I could be doing it wrong and that, you know what I mean, I mightn't be feeling . . . Yeah, I mean I would have liked then to be able to phone up, you know, the physio[therapist] and say 'look, can I have another appointment?', rather than the length of time between each, and then of course it stopped for so many months . . .

Case 6, 12M, basic

Otherwise, the main barriers for women in maintaining PFMT, whether allocated to the biofeedback PFMT or basic PFMT group, were similar to those found in the active treatment phase. First, some women's UI had improved to such an extent that they had no symptoms to act as a reminder to exercise:

... as I said my symptoms have reduced so there's not so much of a physical reminder that 'oh, I need to do them' [PFMT].

Case 28, 12M, biofeedback

The second key barrier was the loss of motivation or loss of the habit of doing PFMT due to life events taking over, even if this was contrary to the intent they had at the end of active treatment. Women spoke of various contextual factors in their lives that prevented them from maintaining a PFMT regime, such as having too many other things to do, work commitments or work changes getting in the way, or more generally feeling that they lacked support. Commonly, women talked of having other non-UI health problems that overshadowed their focus on UI and/or on their attention being more on the needs of others (commonly immediate family). In keeping with the active treatment phase, the findings suggest an interaction between a lack of time (e.g. as shown below, women not having time for themselves) and the multiple other contextual factors that get in the way of life:

Well I've had a lot o' other health issues so it's kinda been, that's [PFMT] been the least o' my worries [UI] tae be honest wi' yae [slight laugh].

Case 10, 24M, basic

Case 32, 12M, biofeedback: ... it really is down to me ... I expect you hear that from a lot of women ... And it's very hard to put yourself at the top of your own time agenda ...

Researcher: As women ...

Case 32, 12M, biofeedback: Yeah, yeah [talks about her husband exercising every day no matter what] ... So, but with me something seems to come up, [then it's?] all my stuff goes to pot on my own agenda ... I suppose it's just, [he's] not easily distracted but there are more pressures on me ... and I think it's probably the same for women generally.

Other factors that acted as barriers to adherence in the maintenance phase were as follows:

- Not establishing PFMT as part of life. Some women found maintaining a PFMT exercise programme to be difficult because they had no routine in life generally, or for PFMT specifically, or that routine changed (e.g. going on holiday).
- Not feeling confident in their PFMT technique and when to use it after treatment had stopped. This seemed to manifest as a lack of confidence in (1) their ability to undertake PFMT generally or (2) how to get restarted after a break in PFMT. Various reasons can be identified for this lack of confidence in the maintenance phase: forgetting what they were taught, not feeling that PFMT was going to work, not perceiving that their UI was caused by pelvic floor weakness (it was caused by something else) or because they had not seen symptomatic improvement during active treatment. However, there was a stronger pattern for women to feel confident about continuing PFMT than not having the confidence to continue.
- Ownership and agency in PFMT. Some women talked about a lack of motivation and willpower, they talked about forgetting to exercise (sometimes or always), and PFMT lost the novelty factor and priority over time.

In summary, large-scale systematic differences between the biofeedback PFMT and basic PFMT groups in barriers to PFMT maintenance were not evident from the data set. Key barriers to maintaining PFMT lay in loss of support following the active treatment phase and busy lives.

Women's urinary incontinence outcomes in the short and long term

The case study did not set out to explore outcome, but women discussed outcome as part of their experience. Given the longitudinal case study design, and the core aim of the trial, it was useful to consider women's views of outcome in this chapter in relation to UI symptoms. However, interviewed women reported outcomes that were considerably broader than UI symptoms alone. For example, the women talked about changes they made to their lifestyle, changes to their feelings about UI and about a myriad of things they had learned from being part of the trial. These additional outcomes will be documented in more detail in future publications.

At 24 months (when the primary outcome was measured in the trial) there was no obvious difference between the groups in UI severity from qualitative comparison; rather, there were women in both groups with varying outcomes (Table 33).

In both the biofeedback PFMT and basic PFMT groups there were more women talking about positive outcomes in relation to their UI symptoms at 24 months than there were talking about poor outcomes (i.e. from baseline it seemed as if women tended to be better than they were before they entered the trial). This information, however, needs to be considered with caution, as qualitative studies do not aim to statistically generalise:

I was just going to say well no, thank you for the opportunity because I've seen a massive, you know, improvement and because I've got a prolapse and obviously, I'm quite young, I'm only 38, it was making me sort of anxious about [?] and you know, everything has improved, my bladder control and my prolapse symptoms have improved, I'm not getting as many em, I used to get sort of quite a lot of dragging sort of tummy ache [muscle, or little,?] and I don't get that any more, so, you know, and I know that that is definitely all down to the trial, if I wouldn't have been involved in that, then I know that I'd still be having the problems and still be anxious, you know, if I went out walking or if I went, went running, or to the gym or whatever, so, so I'd like to say thank you to you guys as well.

Case 28, 24M, biofeedback

In terms of short-term outcomes, in both groups there was a pattern that suggested that women were likely to have better UI outcomes at 6 months (immediately post-active treatment phase) than at 24 months. For example, case 13 (biofeedback PFMT) reported symptomatic improvement at 6 and 12 months, but at 24 months reported that her symptoms were the same or a little worse than when she started the trial. There was, however, variance between individuals. For example, for case 32 (biofeedback PFMT), there was no improvement noted at 6 months and at 24 months her symptoms were worse than when she started the trial. There were other cases when improvements occurred beyond 6 months (i.e. 6 months was not the best outcome point). For example, case 36 (basic PFMT) reported good improvement at 6 months, further improvement at 12 months and yet further improvement at 24 months.

TABLE 33 Case study examples of variance in UI outcomes at 24 months by allocated treatment group

Nature of outcome	Treatment group	
	Biofeedback PFMT	Basic PFMT
Good outcome	Case 27 was almost cured with few SUI symptoms at 24 months, some occasional urgency persisted	Case 20 was almost cured; her stress symptoms were gone completely and urgency occurred only occasionally
Intermediate outcome	Case 17's symptoms were not gone but were much improved (e.g. she makes it to the toilet with UI most of the time)	Case 36 continued to have UI symptoms, but they were better than before she started with the OPAL trial (e.g. she had more time to get to the toilet)
Poor outcome	Case 32's UI symptoms were worse at 24 months than when she started in the trial	Case 24's symptoms were the same or worse at 24 months than when she started in the trial

In summary, there were no obvious differences in UI outcome between the trial groups.

Theoretical propositions

Two theoretical propositions and one rival explanation were considered. The theoretical propositions were driven by the theory that supported the hypothesised mechanism of action (propositions 1 and 2) and one rival explanation that arose from analysis of the data (proposition 3).

Proposition 1: biofeedback PFMT will improve (1) women's adherence and (2) women's urinary incontinence outcomes more than basic PFMT in the short and long term

This proposition was the main hypothesis of the trial. There was no clear evidence that biofeedback PFMT improved adherence over basic PFMT in the short or long term, nor any clear evidence of greater improvement of outcomes in either the short or the long term. Therefore, the theoretical proposition was not supported.

Proposition 2: the factors that influence women's adherence and women's urinary incontinence outcomes change over time

This proposition arose from the long-term nature of the follow-up that was part of the commissioning brief and was based in our understanding of the influence of context (e.g. Wells *et al.*⁶⁷). This proposition was supported in that it was clear that the factors that influence adherence and outcome for an individual woman do change over time. For example, there were women who were diagnosed with other conditions during the trial that, for them, took precedence in their quest for good health. However, the hypothesis aimed to identify if there were factors that arose at specific time points for a group of women. It does not seem that there were factors that occurred at the same time point in specific groups of women (other than the removal of support when treatment finished); however, this will be the subject of further analysis.

Proposition 3: factors other than biofeedback PFMT or basic PFMT will influence adherence and urinary incontinence outcome in the short and long term

This proposition arose from rival explanations (to biofeedback PFMT or basic PFMT directly linking to adherence and outcome) being identified iteratively in data analysis. Although there were factors other than the interventions that influenced adherence and outcome, there were considerably more similarities in the factors than differences between the groups. The notion of life events taking over encapsulates this well. However, for some women with multiple other life events, there was still adherence and a symptomatic improvement (i.e. these factors did not always act to diminish adherence or outcome, but they often did).

Discussion

Women reported positive experiences of both the biofeedback PFMT and basic PFMT interventions; in particular, women were clear about the benefit of therapist input. There were no major differences, based on qualitative comparison, in adherence to PFMT or UI outcome between the biofeedback PFMT and basic PFMT groups, with wide variation in adherence and outcome in both groups. Adherence in the short and long term was facilitated by women's desire to improve or cure their UI symptoms and by factors related to the therapists, which included feedback given through vaginal examination and rapport. A lack of time and life taking over were key barriers to adherence in both the short and long term. Adherence did change over time, but there were no clear differences between the groups. Although UI outcome did not appear to differ between the groups, there was a trend towards improved outcomes at 2 years when compared with baseline. There were features of biofeedback PFMT that worked as anticipated (such as visualisation), but there were also drawbacks to biofeedback (such as it taking more time than PFMT alone).

Strengths and limitations of the case study

A key strength of the case study, and qualitative research linked to trials in general, is that it facilitated the voice of those whom the intervention aimed to help to be heard and represented. The longitudinal nature of the case study, with detailed follow-up at the same time points as the trial, and the purposeful searching for expansion on emerging ideas at subsequent interviews, allowed consideration of women's expressions

of adherence to PFMT over time. Studies of long-term adherence in UI are rare (only one other longitudinal study,⁶⁸ with women who have UI, has been identified), but are important as reduced adherence is a common explanation for why treatment effect is not sustained over time.⁶⁹ The two-tailed case study design offers a robust, qualitative means of comparison that supports the comparison in the trial.

The process evaluation and case study drew on a contemporary published framework in further developing the work and developing the analysis plan.⁶³ That framework proposes multiple candidate approaches to understanding various features of the trial and its effects. However, one weakness was that data were not gathered on all the candidate approaches.⁶³ However, data were gathered on several candidate approaches that were central to the research questions, such as maintenance. Another potential weakness of the case study in relation to the trial is that the interviews may have acted as a co-intervention to promote adherence, for example women reflected that they undertook PFMT because they knew that an interview was coming up. However, the case study recruited women from both the biofeedback PFMT and basic PFMT groups and any effect of the interviews on adherence potentially occurred equally in both groups.

Comparison of findings to existing literature

The evidence from the case study is consistent with the trial finding that biofeedback PFMT did not improve UI outcomes more than basic PFMT. Insights from the case study are helpful in explaining the main trial finding. The qualitative data demonstrated that biofeedback could work as anticipated, with women reporting the benefits of being able to visualise the contraction and know that they were doing the contraction correctly, alongside their learning in partnership with the therapist. However, women in the basic PFMT group also had confidence in their ability to undertake PFMT. For these women, this was based on learning in partnership with the therapist. A possible conclusion is, therefore, that biofeedback does not need to be added to a strong basic PFMT programme in order for women to achieve self-efficacy for and adherence to PFMT; good therapist input can also provide self-efficacy and adherence.

Aspects that are central to this conclusion are that the OPAL trial basic PFMT (and biofeedback PFMT) programmes allowed sufficient time with therapists to support a treatment effect;⁷⁰ both interventions were based on BCTs;⁸ both demonstrated that some women achieved self-efficacy for PFMT;⁷¹ both groups received therapist-mediated vaginal feedback; and one group also received biofeedback.⁷ Although it is possible that if biofeedback PFMT had been compared with a less robust basic PFMT programme there would have been a difference between groups, it would then have been difficult to reach conclusions about the effectiveness of adding biofeedback because of other confounding variables.⁷ Our conclusions therefore support a finding that, if all other aspects of PFMT are kept equal, the addition of biofeedback may not lead to a greater improvement in continence outcomes.

Another possible explanation for why biofeedback PFMT was not more effective than the basic PFMT is that, although women in the biofeedback PFMT group did identify features of biofeedback as facilitators of adherence, they also identified features of biofeedback as barriers to adherence. One tentative hypothesis here is that the facilitators and barriers simply cancel one another out. However, this needs further analysis.

Case study findings demonstrated that women generally reported positive experiences of the OPAL trial interventions. Women were positive about learning to do PFMT (with or without biofeedback), which is consistent with a previous qualitative synthesis.¹⁸ Women were also very positive about the therapists. Women talked about the therapist in ways that suggested that the therapist was seen as a credible source of information, a motivator and as someone who could support the learning of the necessary behavioural skills, all of which supports the theory underlying the development of the interventions (IMB¹⁷). Furthermore, in keeping with previous suggestions, rapport between woman and therapist was seen as a factor in supporting adherence.¹⁸

There was a trend identified in the case study data for women to perceive that their UI was better at 2 years than it was when they started the trial. This was not the case for all women. It was, however,

an important finding in the context of the worldwide evidence that UI negatively affects women's day-to-day lives (see, for example, Bradway,⁷² Delarmelindo Rde *et al.*⁷³ and Hamid *et al.*⁷⁴). Although it is possible that the improvement described by women is not linked to the interventions, the evidence suggests that women did perceive a link between the intervention they received, their adherence to PFMT and their positive outcome. This link will be explored in more detail in further analysis.

Adherence to PFMT did change over time, but not differently between the allocated groups. A key reason for including the case study alongside the trial was the recognition of the influence of context on the effectiveness of complex interventions.^{67,75} It is now widely recognised that context interacts, modifies, shapes and constrains the intervention and implementation.⁷⁶ This study chose to investigate the influence of context in-depth from the participants' perspectives (rather than also exploring the problem, trial and organisational contexts), because it was believed these would be the most important factors to shape the interventions and influence their effectiveness. It is important to understand the dynamic relationship between context, implementation and intervention to define what was implemented and understand how works in certain contexts. It is now no longer enough to say what works, we need to explore what works, for whom and in what context.⁵³ It was clear that many varied personal contextual factors influenced adherence. The longitudinal nature of the study was important in highlighting that for all women, context and implementation were dynamic and life events got in the way. In addition, many women put the needs of others before themselves. We need to carry out a nuanced analysis to explore the characteristics of these women to understand the various ways women may overcome these events. Previous research supports the links between life taking over and inconsistent adherence in UI.^{57,77} These findings suggest that when delivering a PFMT intervention, or in future research, consideration should be given to helping women balance the multiple contextual factors in ways that may support their engagement with PFMT and their re-engagement with PFMT after a break.

Urinary incontinence symptoms were an important factor in adherence at the outset, and continued to be so in the long term. Symptoms influenced adherence in a number of ways. Women adhered to rid themselves of symptoms, but, conversely, when symptoms were no longer present, the trigger to exercise was no longer there and some women then stopped exercising. Women had to perceive change and believe that it was linked to treatment to maintain adherence in the longer term. This finding is consistent with other studies.¹⁸ It is an important feature of care delivery for therapists to keep the connection between PFMT and symptomatic improvement at the forefront of women's minds.

Chapter 7 Results synthesis

Aims

The OPAL trial consisted of three main components: a trial, a process evaluation and a longitudinal case study. The aim of the synthesis was to bring together findings from individual data sources to facilitate drawing overall conclusions about why the interventions may or may not have been clinically effective [see the project web page: www.journalslibrary.nihr.ac.uk/programmes/hta/117103/#/ (accessed 29 July 2019)] and identify where there was agreement and disagreement between the findings.

Specifically, agreement and disagreement was assessed in the four following key areas:

1. benefit of biofeedback PFMT over basic PFMT in UI symptoms
2. whether or not the biofeedback PFMT group received an intensified intervention that could have led to improved continence outcomes
3. whether or not the biofeedback PFMT group had greater self-efficacy for PFMT that could have encouraged more PFMT and therefore improved continence outcomes
4. whether or not the biofeedback PFMT group had greater adherence to PFMT that could have led to improved continence outcomes.

Methods

To achieve this aim, an approach to integrating of results from mixed-methods research described⁷⁸ was adopted: a triangulation protocol.⁷⁸ The term triangulation is used here to mean a way of addressing a study question using different methodologies to give a fuller picture. Triangulation occurs at the interpretation stage, once the contributing data sets have each been analysed separately. One form of triangulation protocol involves producing a convergence coding matrix, which shows findings from each of the individual study components, with an assessment of whether there is agreement, partial agreement, or disagreement between them. Occurrences of 'silence' are also noted when there is no information from one source on a theme arising in another.

We applied this approach and produced a convergence coding matrix based on the five data sources available:

1. data relating to all trial participants, recorded in questionnaires at each time point, exercise diaries and TAFs (trial)
2. data from interviews with a sample of trial participants (case study)
3. data from interviews with therapists as part of the process evaluation
4. data from the checklists completed by therapists as part of the TAF at every appointment as part of the process evaluation (checklists)
5. data from audio-recordings of a sample of therapy appointments as part of the process evaluation (audios).

Results

The results of the synthesis are summarised in *Table 34*. There was evidence of full agreement from all data sources that there was no group difference in severity of UI at the 24-month follow-up. The primary outcome of UI severity measured by the ICIQ-UI SF at 24 months demonstrated no group difference in UI

TABLE 34 Convergence coding matrix

Finding	Trial	Case study	Therapist interviews	Checklists	Audios	Agreement
No benefit of biofeedback PFMT over basic PFMT	<p>ICIQ-UI SF score at 24 months did not differ significantly between groups, indicating that biofeedback PFMT did not result in less severe UI than basic PFMT</p> <p>There was no difference between groups at 24 months when expressed as cure or improvement in UI</p>	<p>Women in both groups reported a range of good, intermediate and poor UI outcomes</p> <p>At 24 months there was no obvious group difference in changes in UI severity from baseline</p> <p>Some women in each group valued improvement in UI symptoms, even if this was not what might be measured as cure</p>	<p>Some therapists expected there to be no difference between groups</p> <p>Barriers to therapist's delivery and women's receipt of biofeedback may have had an impact on the potential benefits</p> <p>Simplicity of delivery of the basic PFMT protocol and of implementing it at home may have given this group an advantage</p>	Silence	Silence	Full
Biofeedback PFMT group received an intensified intervention	<p>In the intervention phase, in the biofeedback PFMT group:</p> <ul style="list-style-type: none"> 80% received biofeedback during one or more appointments 78% used the biofeedback unit one or more times at home 	<p>Silence (further post hoc analysis will explore relevant findings from the case study)</p>	<p>Therapists liked the biofeedback PFMT protocol, but felt that they were pressed for time</p> <p>Therapists experienced in using biofeedback said that in their usual practice, they would not use biofeedback in the first appointment</p>	<p>Therapists did deliver biofeedback PFMT and women did receive it; the biofeedback PFMT group did get more overall</p>	<p>The biofeedback group got more, but some core basic PFMT components may have been omitted as a result in some sessions (e.g. practising 'the Knack')</p>	Partial

Finding	Trial	Case study	Therapist interviews	Checklists	Audios	Agreement
Biofeedback PFMT group had greater self-efficacy for PFME contraction	<p>PFME self-efficacy score at 24 months was significantly better in the biofeedback PFMT group vs. basic PFMT group</p> <p>The addition of biofeedback may be associated with women feeling more confident about the contractions and exercises, in the context of their lives and in gaining benefits</p> <p>However, this may not be a clinically meaningful difference (only 2.4-point difference)</p>	<p>Some biofeedback PFMT women reported positives of having biofeedback in line with the hypothesised mechanism of action (ability to visualise the contraction giving confidence in ability to contract correctly)</p> <p>Women in the basic PFMT group also reported confidence in their ability to undertake a contraction and the exercises</p> <p>At the current stage of analysis, there is no obvious difference in self-efficacy between groups. Some women in each group reported confidence in their ability to undertake PFMT at 24 months</p>	<p>Some therapists commented that biofeedback empowered women</p> <p>All therapists said that in practice, they would use biofeedback for women who struggled to feel their pelvic floor muscles</p> <p>Monitoring of contraction strength/repetitions with biofeedback was thought to increase confidence and women's accountability to engage/attend appointments</p> <p>Goal-orientated women were thought, perhaps, to be most suited to biofeedback PFMT</p>	Silence	Silence	Partial
Biofeedback PFMT group had greater adherence to PFMT	<p>No difference in adherence between groups in terms of the number of appointments attended</p> <p>Engagement at home:</p> <ul style="list-style-type: none"> similar percentage adhering to home programme during intervention phase (78% biofeedback vs. 81% basic) similar percentage exercising a few time per week or more at 24 months (49% biofeedback vs. 43% basic) 	<p>Women in both groups reported varying levels of adherence, with no obvious group differences in the ways women adhered in terms of home exercise in the shorter or longer term</p> <p>There were many barriers and facilitators that influenced adherence in both groups, with more similarity between the groups than difference</p>	<p>Some therapists thought that some women adhered better because they had biofeedback – if women were goal-orientated, or had the privacy and time to use at home</p> <p>'Buy-in' to the intervention was reported to matter and may have affected adherence (e.g. if women expected biofeedback to stimulate their muscles, they were disappointed and became demotivated)</p>	Silence	Silence	Partial

severity; this finding held when different assumptions were tested. The case study also showed that, across the time points, women in both groups reported varying degrees of improvement (see *Table 34*). The therapist interviews highlighted some issues with using biofeedback in clinic, which may have affected delivery of the protocol. The case study also made reference to difficulties with biofeedback being experienced by women. Biofeedback was, however, reported as being delivered in 80% of appointments and used at home by 78% of participants in this group. The addition of biofeedback did not improve continence outcomes at 2 years and a robust basic PFMT protocol was shown to be as effective as PFMT plus biofeedback.

There was a conclusion of partial agreement regarding whether or not the biofeedback PFMT group unequivocally received a more intensive intervention. There was evidence from the checklists that participants in the biofeedback PFMT group did receive biofeedback in clinic and used biofeedback at home, which the basic PFMT group did not have access to. From the checklists, it also appeared that the biofeedback PFMT group received all elements of the protocol, as did the basic PFMT group (i.e. the only difference between groups was the addition of the clinic and home-based biofeedback). There was no evidence from our preliminary analysis of the interviews with women in the case study to allow comment on whether the biofeedback PFMT group did or did not receive the intended intensified protocol. A preliminary analysis of the therapist interviews and appointment audio-recordings might suggest that some elements of the basic PFMT protocol had been left out, or were less well emphasised, in order to accommodate delivery of biofeedback in clinic. A further analysis is planned to investigate this by doing a more detailed cross-checking of the data sets.

Greater self-efficacy was hypothesised to be a mechanism by which biofeedback would increase the quality and quantity of PFMT that women undertook, leading to improved continence outcomes. There was partial agreement for this hypothesis from the data sources. The participants' responses to the self-efficacy for PFME scale³⁴ indicated a small significant difference between the trial groups, with the biofeedback PFMT group having greater self-efficacy. There is no published minimal clinically important difference for this scale to allow assessment of the probable importance of the observed difference. However, it seems improbable that a 2.4-point difference on a scale ranging from 17 to 85 (e.g. equating to an improvement in one item from 'strongly disagree' to 'neutral') would be meaningful. Interviews with women provided evidence of self-efficacy in both groups, whereas therapists saw the potential for biofeedback to further empower some women.

There was partial agreement that initial (uptake and adoption) and longer-term (maintenance) adherence was no greater in the biofeedback PFMT group in terms of how well women engaged with the interventions. In the main trial, adherence was seen to be similar between the groups: similar numbers of appointments were attended and home exercise was comparable over time. Similarly, no overall difference in adherence between groups emerged from the interviews with women, with many barriers and facilitators being the same for both groups. However, interviews with both women and therapists indicated that there were also differences between the groups: some women did adhere better as a result of clinic biofeedback, and some women lacked time and privacy to fit in home biofeedback.

Discussion

The synthesis of the available data is a work in progress and conclusions here are tentative; however, on the whole, there was partial, if not complete, support for the main trial findings from the study components designed to offer deeper understanding of the complex interactions and main 'no difference' findings. We hypothesised that biofeedback would facilitate self-efficacy, which in turn would increase adherence to PFMT and better outcome; however, none of these have been upheld when we have brought together the evidence and triangulated four of the main theoretical components and associated outcomes.

Chapter 8 Discussion and conclusions

Statement of principal findings

The results of the OPAL trial showed clearly no evidence of a difference between biofeedback PFMT and basic PFMT in terms of UI severity for women with SUI or MUI at 2 years. This was supported when we examined all urinary outcomes (cure/improvement, other lower urinary tract symptoms, condition-specific quality of life, patient perception of UI improvement) both at the 2-year time point and the interim 6- and 12-month time points. Further evidence of no difference between groups was found in the other secondary outcome measures of pelvic floor function, prolapse symptoms, bowel symptoms and uptake of other UI treatment. There was a small difference in self-efficacy between groups, perhaps indicating a trend towards the biofeedback PFMT group having more confidence relating to PFMEs, in line with the hypothesised effect of biofeedback, but the magnitude of the difference was small and unlikely to be clinically significant.

Adherence to the intervention protocols, a potential mediator of better clinical outcomes, was similar between groups, again in keeping with the main trial finding. There were no SAEs related to the intervention and only one non-serious related AE, which was found to be due to a nickel allergy in a biofeedback PFMT participant.

Both interventions were found to be similarly cost-effective. The process evaluation highlighted some challenges with, and barriers to, biofeedback delivery from the therapists' perspective on one hand, but offered evidence of fidelity to the protocol nonetheless. Interviews with participants in the case study supported the finding of no evidence of a difference between groups in the UI outcome.

Although there was no evidence of a difference between the two randomised groups, improvement in ICIQ-UI SF was observed in both treatment groups over the course of the trial, with 8% of women reporting cure and 60% improvement in each group at 24 months.

Robustness of primary outcome analysis

The OPAL trial was large enough to detect a meaningful difference between groups if one existed, with a sample size far exceeding the largest previous trial of its type ($n = 238$, PFMT plus biofeedback and biofeedback vs. PFMT alone).⁷⁹ There was limited information about the minimal clinically important difference of the ICIQ-UI SF on which to base the sample size calculation at the outset of this trial; however, a recent publication indicated that a 3-point difference was clinically important, supporting our original assumption.²⁸ In addition, the observed SD of the primary outcome was 5 points, which was lower than our original assumption of 10 points, again pointing to the trial having sufficient statistical power.

In sensitivity analyses we showed that our primary outcome analysis was robust to the effects of missing data (tested under assumptions of missingness at random and not at random) and treatment non-compliance.

Further analyses of the primary outcome

In further analysis of the primary outcome, we also examined whether or not specific subgroups benefited from biofeedback PFMT compared with basic PFMT. The mechanism for the effect of PFMT in urgency UI differs from that for SUI; therefore, we investigated whether or not there was a different treatment effect for those women with pure SUI and those with a mix of SUI and urgency UI. There may be reasons why older, more so than younger, women might do less well with the addition of biofeedback, such as older women being less likely to engage with the biofeedback technology. The varying levels of severity of UI might be due to different underlying problems, which could respond differently to the addition of biofeedback. Finally, we explored whether or not women treated by different types of therapists might

have different responses to the interventions. There was no evidence to suggest that women who differed with respect to type of UI (SUI vs. MUI), age (< 50 years vs. \geq 50 years), UI severity (mild/moderate vs. severe) or therapist type (physiotherapist vs. non-physiotherapist) differentially benefited from either intervention type.

Strength and limitations

Strengths

A major strength of the OPAL trial was its mixed-methods design, including a RCT, with a process evaluation and longitudinal case study alongside. This offered the potential for a much deeper understanding of the intervention mechanisms and trial results from different perspectives.

The trial itself was conducted and reported following CONSORT recommendations [URL: www.consort-statement.org/ (accessed 31 July 2019)]. We used an automated computer randomisation application developed by CHaRT, accessed remotely by centre staff, to assign group allocation, thus minimising the risk of selection bias. The analyses were conducted on an ITT basis and in accordance with a predefined statistical analysis plan [see the project web page: www.journalslibrary.nihr.ac.uk/programmes/hta/117103/#/ (accessed 29 July 2019)], agreed with the TSC. We chose a well-validated and commonly used instrument as our primary outcome measure (ICIQ-UI SF) and a long duration of follow-up.

Participants were recruited from various outpatient and community settings, adding to the generalisability of the findings. At baseline, the two groups were highly comparable, indicating that the allocation process implemented in the trial had been successful. Our follow-up questionnaire response rates were high, indicating a low risk of attrition bias, with rates of 74% at 6 months, 84% at 12 months and 78% at 24 months. The lower response rate at 6 months is probably because of a less intensive process for contacting non-responders at this time point, which was not the primary end point. Such high retention rates limited the impact of missing data on our findings and this was confirmed by the sensitivity analyses, which showed no impact on the primary outcome analysis findings of various assumptions about missing responses.

All elements of the trial delivery were standardised, as far as possible, to minimise the risk of performance bias. Therapists in trial centres received training in all aspect of the interventions, including using the biofeedback equipment, delivering the PFMT protocol and incorporating the core and optional BCTs. This information was also provided to all therapists in a detailed intervention manual.

Limitations

Reasons for exclusion of women from the trial were kept to a minimum; however, one reason for women being ineligible was being unable to contract their pelvic floor muscles. Such women were excluded based on the NICE guideline, which recommends that women who cannot actively contract their pelvic floor muscles should be offered biofeedback.¹ If we had included these women, half would have been allocated to basic PFMT and therefore denied biofeedback. This means, however, that conclusions cannot be drawn from the trial about this subpopulation of women with UI.

It was not possible to blind participants as to which group they were in because of the nature of the interventions, and for the same reason it was not possible to blind treating therapists. The risk of detection bias was, therefore, high, given that participants were the main source of outcome data. The pelvic floor assessment at 6 months was, however, carried out by a clinician who was blind to the participant's group and, in keeping with other outcomes, the finding was no evidence of a difference between groups.

Approximately one-third of women in both groups attended all six of their scheduled therapy appointments (109/295, biofeedback PFMT and 106/298, basic PFMT). Women commented in interviews on the lack of time to attend frequent appointments. This would reflect the situation in clinical practice, in which the 'did not attend' rate for PFMT appointments is high.

Another potential limitation is that biofeedback equipment is advancing rapidly and biofeedback units are now available that incorporate bluetooth and smartphone technology. These new features, not widely available at the time of developing this trial, may mean that therapists and women would now be less constrained in using such devices. Thus, there may be potential for an enhanced effect of biofeedback delivered using newer devices. These advancements may also have overcome some of the technical issues associated with biofeedback reported by therapists and women in the OPAL trial.

Interpretation of results

Chapters 4–6 provided discussions of the health economics, process evaluation and case study findings, respectively. Here, we consider how to interpret the totality of the OPAL trial findings, drawing on the synthesis described in *Chapter 7*.

Our main research objective in the OPAL trial was to compare biofeedback PFMT with basic PFMT in terms of clinical effectiveness and cost-effectiveness. We met this objective and concluded that there was no evidence of a difference in clinical outcomes and there was no difference in cost-effectiveness. The process evaluation, with its focus on intervention fidelity, confirmed that an intensified PFMT intervention was delivered in the biofeedback PFMT group as intended, but this did not lead to improved clinical outcomes for participants in this group. The case study focused mainly on women's adherence and found varying levels of adherence in both groups, with the barriers to and facilitators of adherence (such as self-efficacy) having more similarities than differences. The finer-grain analysis described in *Chapters 5 and 6* found some possible reasons why there were no differences between trial groups. These included difficulties associated with the delivery of biofeedback in clinic, and the influence of other contextual and time factors influencing the opportunity for women to use biofeedback (or undertake PFMT) at home. The interpretation of the study findings, however, remains, that a benefit of biofeedback was not found. This is in keeping with the recommendations in national guidelines, stating that EMG biofeedback should not be used as a routine part of PFMT.¹

The Cochrane review relating to biofeedback as an adjunct to PFMT for UI, published in 2011, concluded that there may be a benefit of adding biofeedback to PFMT, but that the apparent benefit could be due to differences between the trial arms in the amount of health professional contact.⁷ The OPAL trial sought to address this shortcoming in previous trials by specifying the same number and duration of appointments in both trial groups. Despite this, participants in the biofeedback PFMT group received 26 minutes more contact over the duration of their treatment than those in the basic PFMT group. Even with this extra contact time, however, there was no difference between groups.

A number of trials of biofeedback in women have been published since the publication of the Cochrane review. These have been characterised by small sample sizes and short-term follow-up. Four trials^{80–83} were not directly comparable to the OPAL trial. Ibrahim *et al.*⁸⁰ carried out a randomised comparison of manometric biofeedback-assisted PFMT with PFMT alone in a mixed group of 52 women with either SUI ($n = 26$) or faecal incontinence, with or without prolapse ($n = 26$). At 12 weeks, more women in the biofeedback PFMT group improved their symptoms than in the PFMT alone group. Ahadi *et al.*⁸¹ carried out a randomised pilot trial of PFMT compared with PFMT plus biofeedback in 40 women with prolapse, with short-term outcomes suggestive of better quality of life in the biofeedback PFMT group at 12 weeks. Liu *et al.*⁸² carried out a comparison of biofeedback plus electrical stimulation compared with conventional PFMT and no treatment in perimenopausal SUI. In before-and-after comparisons, both groups receiving treatment improved, whereas the no-treatment group did not. Terlikowski *et al.*⁸³ also compared EMG biofeedback plus electrical stimulation with a placebo for SUI ($n = 112$). It was unclear what the placebo was, but this group had significantly worse urine leakage after 8 weeks.

Five trials^{84–88} were more directly comparable to the OPAL trial and had mixed findings. Hirakawa *et al.*⁸⁴ compared PFMT plus biofeedback (unspecified) with PFMT for SUI ($n = 46$) and found no differences between groups after 12 weeks in any of the parameters assessed, including the King's Health Questionnaire. In a similar study, Fitz *et al.*⁸⁵ also compared PFMT plus biofeedback (unspecified) with PFMT for SUI ($n = 40$) and did find differences between groups in the King's Health Questionnaire at 12 weeks, favouring the biofeedback PFMT group in all but one domain. In another trial by Fitz *et al.*,⁸⁶ of 72 women with SUI, PFMT (clinic and home) was compared with PFMT at home augmented with pressure biofeedback in clinic, with a primary outcome of frequency of exercises after 3 months. Biofeedback did not increase the frequency of women's home exercise. Bertotto *et al.*⁸⁷ compared the effect on muscle strength of PFMT with and without EMG biofeedback in women with SUI ($n = 49$). The biofeedback PFMT group had significantly more improvement in muscle strength than the PFMT-only group. Özlü *et al.*⁸⁸ conducted a three-arm trial in women with SUI, comparing PFMT with PFMT plus pressure biofeedback with PFMT plus EMG biofeedback ($n = 53$). Severity of UI, cure/improvement and pelvic floor muscle strength outcomes were superior in both biofeedback PFMT groups compared with PFMT alone.

We plan to pool the findings of our much larger trial results, where possible, with the results of other comparable trial in a meta-analysis. The overall finding is highly likely to be one of no evidence of an effect of adjunctive biofeedback.

Despite the finding of no evidence of a difference in relation to our primary outcome, we did observe an improvement in ICIQ-UI SF scores in both groups. This could suggest that both interventions had a benefit. This is in keeping with current Grade A evidence for PFMT for the treatment of SUI.⁸⁹ The magnitude of the improvement in the scores we observed was in line with the benefit in ICIQ-UI SF reported in the recently updated Cochrane review,⁹⁰ which compared PFMT with no treatment or inactive control, suggesting that the improvements in ICIQ-UI SF score we observed in both groups were real effects, rather than effects of the women being involved in research. The three relevant trials in the Dumoulin *et al.*⁹⁰ review, however, differed from each other and from the OPAL trial in terms of the PFMT intervention content and duration.

The pattern of a sustained improvement in ICIQ-UI SF score was also observed in other outcomes: for all the outcomes with an observed improvement from baseline to 6 months, this was continued throughout follow-up. This was the case despite the fact that most participants had completed therapy appointments in the active treatment phase at the 6-month time point, following which they could no longer rely on the motivating force of their therapist for continuing their PFMT. Encouragingly, almost half of women (49% and 43% of women in the biofeedback PFMT and basic PFMT groups, respectively) were exercising a few times a week or more at 24 months. There are few PFMT trials with similar long-term follow-up to the OPAL trial for comparison. Only one trial out of 31 in the Dumoulin *et al.*⁹⁰ review had follow-up to 1 year, and this trial also found that the effect of PFMT on UI symptoms and condition-specific quality of life persisted.⁹¹

Implications for health care

- Health-care practitioners and policy-makers can be confident that routine use of EMG biofeedback as part of PFMT in all cases does not confer benefit for women with SUI or MUI in terms of their long-term incontinence severity, and is not more cost-effective than PFMT alone. The combined PFMT and biofeedback intervention evaluated did not improve secondary outcomes over PFMT alone.
- Our theoretically developed interventions relied on women carrying out PFMT (and using biofeedback) at home, as well as at appointments. The interventions incorporated a range of established and effective BCTs with which women and therapists engaged. These techniques could be highlighted to women and health-care professionals as potential ways of helping increase adherence to PFMT.

Future research implications

- Future research would be useful to investigate other ways of intensifying PFMT that may improve women's outcomes compared with a basic PFMT programme. Biofeedback was one possible approach, but others exist. Additional health professional contact to support adherence to PFMT is one option, but careful consideration would need to be given to design of such a service and the availability of staff to resource it.
- Although the OPAL trial found no evidence of benefit of biofeedback, if biofeedback devices evolve sufficiently, more research may ultimately be needed to evaluate new technology that would overcome barriers associated with earlier devices.

Acknowledgements

We would like to thank the external members of the TSC for their advice and support for the project: Mr Chris Mayne (University Hospitals of Leicester NHS Trust), Dr Thomas Chadwick (University of Newcastle), Ms Teresa Cook (Independent Consultant, Women's Health Physiotherapist) and Ms Veronica Haggart (Homerton University Hospital NHS Foundation Trust). Our thanks go also to the DMEC, comprising Dr Steff Lewis (University of Edinburgh), Dr Sue Hallam (Tameside and Glossop Integrated Care NHS Foundation Trust) and Mr Simon Emery (Abertawe Bro Morgannwg University Health Board). We are grateful to the participants who supported the study, giving so generously of their time and sharing their experiences with us. Likewise, we are grateful to the therapists and the research and administrative staff at all 23 collaborating centres who provided invaluable assistance to us throughout the study.

Thank you also to Gladys McPherson and Mark Forrest for their contributions to the development and ongoing support of the trial database and for input to the PMG; to Graeme McLennan, CHaRT Director, for his contribution to the PMG; and to Lyndsay Wilson who provided advice from a service user perspective on the preparation of the protocol and on the PMG. We would also like to thank Lorna Kerr and Kim Stewart who provided administrative and budget support for the trial.

Support

Glasgow Caledonian University agreed to act as sponsor for the research and the study was adopted by the NIHR Clinical Research Network Portfolio (number 15841). The NIHR Collaboration for Leadership in Applied Health Research and Care South West Peninsula at the Royal Devon and Exeter NHS Foundation Trust also supported Sarah Dean's position at Exeter during this work.

Contributions of authors

Suzanne Hagen (Professor of Health Services Research) had responsibility for the study and report overall, for delivery of the trial component specifically and for synthesis of findings from the various study components; was involved in early conception of the trial; and made intellectual input to the study design.

Carol Bugge (Associate Professor of Nursing) led the qualitative component of the study, and was specifically responsible for delivery and reporting of the case study; was involved in the conceptualisation of the case study and process evaluation components; and carried out analysis and interpretation of process evaluation data and case study data.

Sarah G Dean (Professor of Psychology Applied to Rehabilitation and Health) was involved in the conceptualisation of the case study and process evaluation components; was responsible for developing the trial interventions; was responsible for delivery and reporting of the process evaluation; and carried out analysis and interpretation of process evaluation data and case study data.

Andrew Elders (Senior Statistician) was responsible for delivery and reporting of the statistical analysis.

Jean Hay-Smith (Associate Professor in Rehabilitation) was involved in the conceptualisation of the case study and process evaluation components; was responsible for developing the trial interventions; and carried out analysis and interpretation of process evaluation data and case study data.

Mary Kilonzo (Research Fellow in Health Economics) was responsible for delivery of the health economics component of the study; and supported the analysis and write-up of the cost-effectiveness data set.

Doreen McClurg (Professor of Pelvic Floor Physiotherapy) was involved in early conception of the trial; made intellectual input to the study design; and was responsible for developing the trial interventions.

Mohamed Abdel-Fattah (Clinical Chairperson in Gynaecology) was involved in early conception of the trial and made intellectual input to the study design; and contributed expertise in clinical trials, generally, and in the area of incontinence, throughout.

Wael Agur (Consultant Gynaecologist) was involved in early conception of the trial; made intellectual input to the study design; and was a local principal investigator.

Federico Andreis (Lecturer Statistician) carried out analysis and interpretation of process evaluation data.

Joanne Booth (Professor of Rehabilitation Nursing) was responsible for developing the trial interventions.

Maria Dimitrova (Research Assistant in Health Economics) conducted analysis and write-up of the cost-effectiveness data set.

Nicola Gillespie (Data Co-ordinator) assisted trial data management.

Cathryn Glazener (Emeritus Professor of Health Services Research) was involved in early conception of the trial and made intellectual input to the study design; and contributed expertise in clinical trials, generally, and in the area of incontinence, throughout.

Aileen Grant (Senior Research Fellow) carried out the fieldwork for the case study and process evaluation, and carried out analysis and interpretation of process evaluation data and case study data.

Karen L Guerrero (Consultant Urogynaecologist) was a local principal investigator.

Lorna Henderson (Research Assistant) conducted the trial data analysis.

Marija Kovandzic (Research Fellow) carried out analysis and interpretation of the case study data.

Alison McDonald (Senior Trials Manager) provided senior trial management advice and support to the trial office staff.

John Norrie (Chairperson of Medical Statistics and Clinical Trials) contributed expertise in clinical trials, generally, and in the area of incontinence, throughout.

Nicole Sergenson (Data Co-ordinator) was responsible for trial data management.

Susan Stratton (Trial Manager) was responsible for the day-to-day operationalisation and management of the trial.

Anne Taylor (Lecturer in Qualitative Analysis) carried out the fieldwork for the case study and process evaluation and carried out analysis and interpretation of process evaluation data.

Louise R Williams (Trial Manager, June 2018) was responsible for trial management and drafted report material.

All authors were involved in commenting on, reading and approving the final report.

Publications

Grant A, Dean S, Hay-Smith J, Hagen S, McClurg D, Taylor A, *et al.* Effectiveness and cost-effectiveness randomised controlled trial of basic versus biofeedback-mediated intensive pelvic floor muscle training for female stress or mixed urinary incontinence: protocol for the OPAL (optimising pelvic floor exercises to achieve long-term benefits) trial mixed methods longitudinal qualitative case study and process evaluation. *BMJ Open* 2019;**9**:e024152.

Hagen S, McClurg D, Bugge C, Hay-Smith J, Dean SG, Elders A, *et al.* Effectiveness and cost-effectiveness of basic versus biofeedback-mediated intensive pelvic floor muscle training for female stress or mixed urinary incontinence: protocol for the OPAL randomised trial. *BMJ Open* 2019;**9**:e024153.

Hagen S, Elders A, Stratton S, Sergenson N, Bugge C, Dean S, *et al.* Effectiveness of pelvic floor muscle training with and without electromyographic biofeedback for urinary incontinence in women: multicentre randomised controlled trial. *BMJ* 2020;**371**:m3719.

Data-sharing statement

Any data-sharing requests are to be made to the corresponding author.

Patient data

This work uses data provided by patients and collected by the NHS as part of their care and support. Using patient data is vital to improve health and care for everyone. There is huge potential to make better use of information from people's patient records, to understand more about disease, develop new treatments, monitor safety, and plan NHS services. Patient data should be kept safe and secure, to protect everyone's privacy, and it's important that there are safeguards to make sure that it is stored and used responsibly. Everyone should be able to find out about how patient data are used. #datasaveslives You can find out more about the background to this citation here: <https://understandingpatientdata.org.uk/data-citation>.

References

1. National Institute for Health and Care Excellence (NICE). *Urinary Incontinence: The Management of Urinary Incontinence in Women*. London: NICE; 2013.
2. Earnshaw J, Lewis G. NICE Guide to the methods of technology appraisal: pharmaceutical industry perspective. *PharmacoEconomics* 2008;**26**:725–7. <https://doi.org/10.2165/00019053-200826090-00002>
3. Bo K, Frawley HC, Haylen BT, Abramov Y, Almeida FG, Berghmans B, *et al*. An International Urogynecological Association (IUGA)/International Continence Society (ICS) joint report on the terminology for the conservative and nonpharmacological management of female pelvic floor dysfunction. *Neurourol Urodyn* 2017;**36**:221–44. <https://doi.org/10.1002/nau.23107>
4. Abrams P, Cardozo L, Fall M, Griffiths D, Rosier P, Ulmsten U, *et al*. The standardisation of terminology of lower urinary tract function: report from the Standardisation Sub-committee of the International Continence Society. *Neurourol Urodyn* 2002;**21**:167–78. <https://doi.org/10.1002/nau.10052>
5. Siroky MB. Electromyography of the perineal floor. *Urol Clin North Am* 1996;**23**:299–307. [https://doi.org/10.1016/S0094-0143\(05\)70312-8](https://doi.org/10.1016/S0094-0143(05)70312-8)
6. Lucca JA, Recchiuti SJ. Effect of electromyographic biofeedback on an isometric strengthening program. *Phys Ther* 1983;**63**:200–3. <https://doi.org/10.1093/ptj/63.2.200>
7. Herderschee R, Hay-Smith EJ, Herbison GP, Roovers JP, Heineman MJ. Feedback or biofeedback to augment pelvic floor muscle training for urinary incontinence in women. *Cochrane Database Syst Rev* 2011;**7**:CD009252. <https://doi.org/10.1002/14651858.CD009252>
8. Michie S, Richardson M, Johnston M, Abraham C, Francis J, Hardeman W, *et al*. The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: building an international consensus for the reporting of behavior change interventions. *Ann Behav Med* 2013;**46**:81–95. <https://doi.org/10.1007/s12160-013-9486-6>
9. Alewijnse D, Mesters I, Metsemakers J, Adriaans J, van den Borne B. Predictors of intention to adhere to physiotherapy among women with urinary incontinence. *Health Educ Res* 2001;**16**:173–86. <https://doi.org/10.1093/her/16.2.173>
10. Chen SY, Tzeng YL. Path analysis for adherence to pelvic floor muscle exercise among women with urinary incontinence. *J Nurs Res* 2009;**17**:83–92. <https://doi.org/10.1097/JNR.0b013e3181a53e7e>
11. Alewijnse D, Metsemakers JF, Mesters IE, van den Borne B. Effectiveness of pelvic floor muscle exercise therapy supplemented with a health education program to promote long-term adherence among women with urinary incontinence. *Neurourol Urodyn* 2003;**22**:284–95. <https://doi.org/10.1002/nau.10122>
12. Bø K. Pelvic floor muscle training is effective in treatment of female stress urinary incontinence, but how does it work? *Int Urogynecol J Pelvic Floor Dysfunct* 2004;**15**:76–84. <https://doi.org/10.1007/s00192-004-1125-0>
13. Imamura M, Abrams P, Bain C, Buckley B, Cardozo L, Cody J, *et al*. Systematic review and economic modelling of the effectiveness and cost-effectiveness of non-surgical treatments for women with stress urinary incontinence. *Health Technol Assess* 2010;**14**(40). <https://doi.org/10.3310/hta14400>
14. Kastelein AW, Dicker MFA, Opmeer BC, Angles SS, Raatikainen KE, Alonso JF, *et al*. Innovative treatment modalities for urinary incontinence: a European survey identifying experience and attitude of healthcare providers. *Int Urogynecol J* 2017;**28**:1725–31. <https://doi.org/10.1007/s00192-017-3339-y>

15. Borello-France D, Burgio KL, Goode PS, Markland AD, Kenton K, Balasubramanyam A, Stoddard AM, Urinary Incontinence Treatment Network. Adherence to behavioral interventions for urge incontinence when combined with drug therapy: adherence rates, barriers, and predictors. *Phys Ther* 2010;**90**:1493–505. <https://doi.org/10.2522/ptj.20080387>
16. Craig P, Dieppe P, Macintyre S, Michie S, Nazareth I, Petticrew M, Medical Research Council Guidance. Developing and evaluating complex interventions: the new Medical Research Council guidance. *BMJ* 2008;**337**:a1655. <https://doi.org/10.1136/bmj.a1655>
17. Fisher WA, Fisher JD. Understanding and promoting AIDS preventive behaviour: a conceptual model and educational tools. *Can J Hum Sex* 1992;**1**:99–106.
18. Hay-Smith J, Dean S, Burgio K, McClurg D, Frawley H, Dumoulin C. Pelvic-floor-muscle-training adherence ‘modifiers’: a review of primary qualitative studies-2011 ICS State-of-the-Science Seminar research paper III of IV. *Neurourol Urodyn* 2015;**34**:622–31. <https://doi.org/10.1002/nau.22771>
19. Hagen S, Stark D, Glazener C, Dickson S, Barry S, Elders A, et al. Individualised pelvic floor muscle training in women with pelvic organ prolapse (POPPY): a multicentre randomised controlled trial. *Lancet* 2014;**383**:796–806. [https://doi.org/10.1016/S0140-6736\(13\)61977-7](https://doi.org/10.1016/S0140-6736(13)61977-7)
20. Hagen S, McClurg D, Bugge C, Hay-Smith J, Dean SG, Elders A, et al. Effectiveness and cost-effectiveness of basic versus biofeedback-mediated intensive pelvic floor muscle training for female stress or mixed urinary incontinence: protocol for the OPAL randomised trial. *BMJ Open* 2019;**9**:e024153. <https://doi.org/10.1136/bmjopen-2018-024153>
21. Grant A, Dean S, Hay-Smith J, Hagen S, McClurg D, Taylor A, et al. Effectiveness and cost-effectiveness randomised controlled trial of basic versus biofeedback-mediated intensive pelvic floor muscle training for female stress or mixed urinary incontinence: protocol for the OPAL (optimising pelvic floor exercises to achieve long-term benefits) trial mixed methods longitudinal qualitative case study and process evaluation. *BMJ Open* 2019;**9**:e024152. <https://doi.org/10.1136/bmjopen-2018-024152>
22. Garber CE, Blissmer B, Deschenes MR, Franklin BA, Lamonte MJ, Lee IM, et al. American College of Sports Medicine position stand. Quantity and quality of exercise for developing and maintaining cardiorespiratory, musculoskeletal, and neuromotor fitness in apparently healthy adults: guidance for prescribing exercise. *Med Sci Sports Exerc* 2011;**43**:1334–59. <https://doi.org/10.1249/MSS.0b013e318213f6fb>
23. Laycock J, Jerwood D. Pelvic floor muscle assessment: the PERFECT scheme. *Physiotherapy* 2001;**87**:631–42. [https://doi.org/10.1016/S0031-9406\(05\)61108-X](https://doi.org/10.1016/S0031-9406(05)61108-X)
24. Messelink B, Benson T, Berghmans B, Bø K, Corcos J, Fowler C, et al. Standardization of terminology of pelvic floor muscle function and dysfunction: report from the pelvic floor clinical assessment group of the International Continence Society. *Neurourol Urodyn* 2005;**24**:374–80. <https://doi.org/10.1002/nau.20144>
25. Miller JM, Ashton-Miller JA, DeLancey JO. A pelvic muscle precontraction can reduce cough-related urine loss in selected women with mild SUI. *J Am Geriatr Soc* 1998;**46**:870–4. <https://doi.org/10.1111/j.1532-5415.1998.tb02721.x>
26. Avery K, Donovan J, Peters TJ, Shaw C, Gotoh M, Abrams P. ICIQ: a brief and robust measure for evaluating the symptoms and impact of urinary incontinence. *Neurourol Urodyn* 2004;**23**:322–30. <https://doi.org/10.1002/nau.20041>
27. Klovning A, Avery K, Sandvik H, Hunskaar S. Comparison of two questionnaires for assessing the severity of urinary incontinence: the ICIQ-UI SF versus the incontinence severity index. *Neurourol Urodyn* 2009;**28**:411–15. <https://doi.org/10.1002/nau.20674>

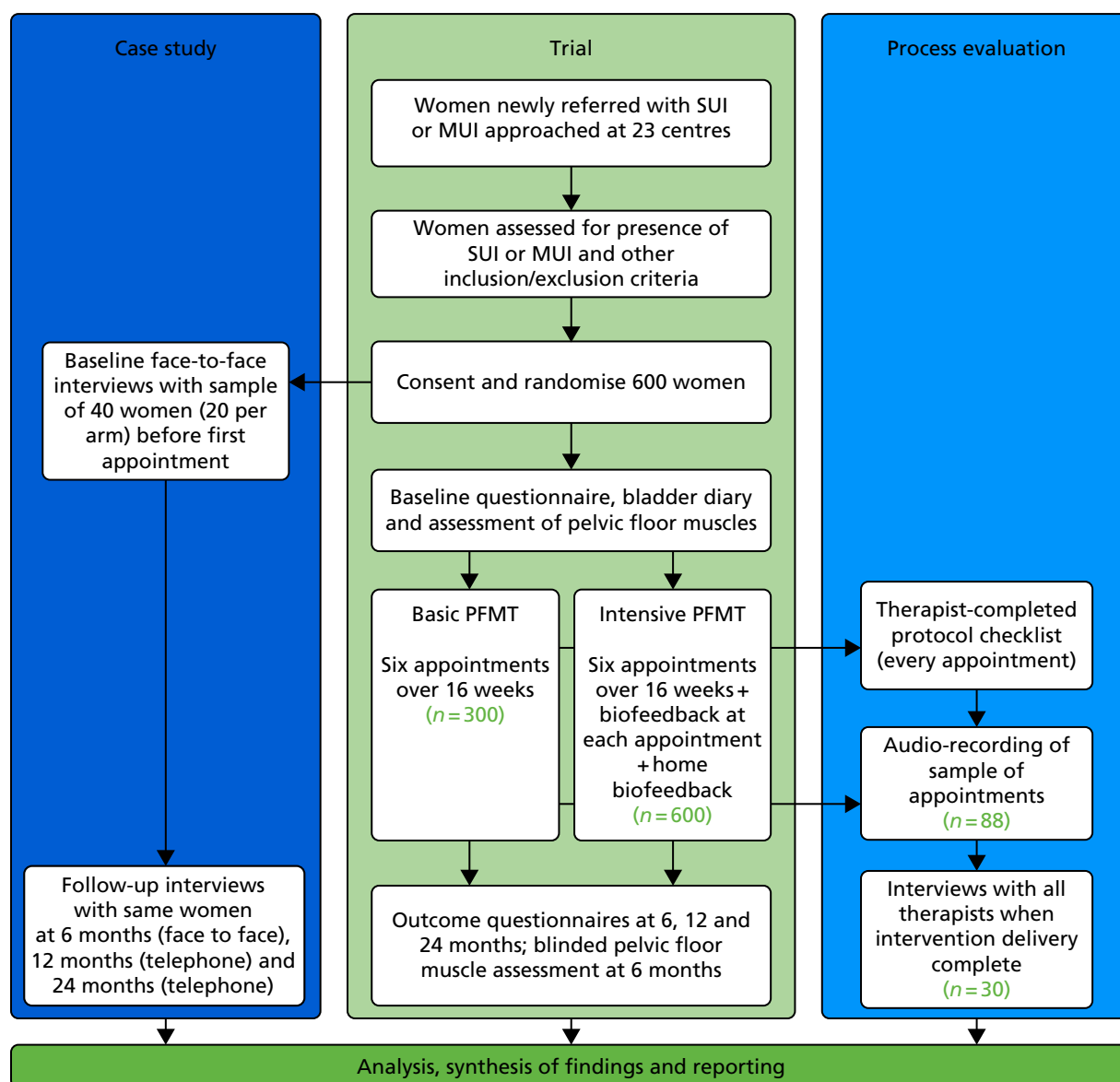
28. Nyström E, Sjöström M, Stenlund H, Samuelsson E. ICIQ symptom and quality of life instruments measure clinically relevant improvements in women with stress urinary incontinence. *Neurourol Urodyn* 2015;**34**:747–51. <https://doi.org/10.1002/nau.22657>
29. Yalcin I, Bump RC. Validation of two global impression questionnaires for incontinence. *Am J Obstet Gynecol* 2003;**189**:98–101. <https://doi.org/10.1067/mob.2003.379>
30. Brookes ST, Donovan JL, Wright M, Jackson S, Abrams P. A scored form of the Bristol Female Lower Urinary Tract Symptoms questionnaire: data from a randomized controlled trial of surgery for women with stress incontinence. *Am J Obstet Gynecol* 2004;**191**:73–82. <https://doi.org/10.1016/j.ajog.2003.12.027>
31. Kelleher CJ, Cardozo LD, Khullar V, Salvatore S. A new questionnaire to assess the quality of life of urinary incontinent women. *Br J Obstet Gynaecol* 1997;**104**:1374–9. <https://doi.org/10.1111/j.1471-0528.1997.tb11006.x>
32. EuroQol Group. EuroQol-a new facility for the measurement of health-related quality of life. *Health Policy* 1990;**16**:199–208. [https://doi.org/10.1016/0168-8510\(90\)90421-9](https://doi.org/10.1016/0168-8510(90)90421-9)
33. Hagen S, Glazener C, Sinclair L, Stark D, Bugge C. Psychometric properties of the pelvic organ prolapse symptom score. *BJOG* 2009;**116**:25–31. <https://doi.org/10.1111/j.1471-0528.2008.01903.x>
34. Chen SY. The development and testing of the pelvic floor muscle exercise self-efficacy scale. *J Nurs Res* 2004;**12**:257–66. <https://doi.org/10.1097/01.JNR.0000387510.52243.c8>
35. Hajebrahimi S, Corcos J, Lemieux MC. International consultation on incontinence questionnaire short form: comparison of physician versus patient completion and immediate and delayed self-administration. *Urology* 2004;**63**:1076–8. <https://doi.org/10.1016/j.urology.2004.01.005>
36. Sherburn M, Bird M, Carey M, Bø K, Galea MP. Incontinence improves in older women after intensive pelvic floor muscle training: an assessor-blinded randomized controlled trial. *Neurourol Urodyn* 2011;**30**:317–24. <https://doi.org/10.1002/nau.20968>
37. Sherburn M, Bø K, Galea M. Evaluation of outcome measures for stress urinary incontinence in older women. *Neurourol Urodyn* 2009;**28**:715–16.
38. Hagen S, Elders A, Stratton S, Sergenson N, Bugge C, Dean S, et al. Effectiveness of pelvic floor muscle training with and without electromyographic biofeedback for urinary incontinence in women: multicentre randomised controlled trial. *BMJ* 2020;**371**:m3719. <https://doi.org/10.1136/bmj.m3719>
39. National Institute for Health and Clinical Excellence (NICE). *The Guidelines Manual. Process and Methods*. 2012. URL: www.nice.org.uk/process/pmg6/resources/the-guidelines-manual-pdf-2007970804933 (accessed 10 June 2018).
40. Curtis L, Burns A. *Unit Costs of Health and Social Care 2017*. 2017. URL: <https://doi.org/10.22024/UniKent/01.02/65559> (accessed 20 July 2018).
41. Joint Formulary Committee. *British National Formulary*. URL: www.medicinescomplete.com (accessed 20 July 2018).
42. NHS Improvement. *Reference Costs 2016/17*. URL: <https://improvement.nhs.uk/resources/reference-costs/#rc1718> (accessed 3 April 2018).
43. Bupa. *Bupa Pay as You Go Healthcare*. URL: www.bupa.co.uk/health/payg (accessed 3 May 2018).
44. Department for Transport. *WebTAG: TAG Data Book, May 2018. Transport Modelling and Appraisal*. London: Department for Transport; 2018.

45. HM Revenue and Customs. *Expenses and Benefits: Business Travel Mileage for Employees' Own Vehicles*. 2018. URL: www.gov.uk/expenses-and-benefits-business-travel-mileage/rules-for-tax (accessed 12 September 2018).
46. Office for National Statistics. *Earning and Hours Worked, All Employees: ASHE Table 1*. 2018. URL: www.ons.gov.uk/employmentandlabourmarket/peopleinwork/earningsandworkinghours/datasets/allemployeesashtable1 (accessed 25 October 2018).
47. Dolan P. Modeling valuations for EuroQol health states. *Med Care* 1997;**35**:1095–108. <https://doi.org/10.1097/00005650-199711000-00002>
48. Brazier J, Czoski-Murray C, Roberts J, Brown M, Symonds T, Kelleher C. Estimation of a preference-based index from a condition-specific measure: the King's Health Questionnaire. *Med Decis Making* 2008;**28**:113–26. <https://doi.org/10.1177/0272989X07301820>
49. Glick H, Doshi J, Sonnad S, Polsky D. *Economic Evaluation in Clinical Trials*. 1st edn. New York, NY: Oxford University Press; 2007. <https://doi.org/10.1093/med/9780199685028.003.0001>
50. Drummond MF, Sculpher MJ, Torrance GW, O'Brien BJ, Stoddart GL. *Methods for the Economic Evaluation of Health Care Programmes*. Vol. 3. Oxford: Oxford University Press; 2005.
51. Barber JA, Thompson SG. Analysis of cost data in randomized trials: an application of the non-parametric bootstrap. *Stat Med* 2000;**19**:3219–36. [https://doi.org/10.1002/1097-0258\(20001215\)19:23%3C3219::AID-SIM623%3E3.0.CO;2-P](https://doi.org/10.1002/1097-0258(20001215)19:23%3C3219::AID-SIM623%3E3.0.CO;2-P)
52. Leurent B, Gomes M, Faria R, Morris S, Grieve R, Carpenter JR. Sensitivity analysis for not-at-random missing data in trial-based cost-effectiveness analysis: a tutorial. *PharmacoEconomics* 2018;**36**:889–901. <https://doi.org/10.1007/s40273-018-0650-5>
53. Moore GF, Audrey S, Barker M, Bond L, Bonell C, Hardeman W, et al. Process evaluation of complex interventions: Medical Research Council guidance. *BMJ* 2015;**350**:h1258. <https://doi.org/10.1136/bmj.h1258>
54. Bandura A. Self-efficacy: toward a unifying theory of behavioral change. *Psychol Rev* 1977;**84**:191–215. <https://doi.org/10.1037/0033-295X.84.2.191>
55. Hawe P, Shiell A, Riley T. Complex interventions: how 'out of control' can a randomised controlled trial be? *BMJ* 2004;**328**:1561–3. <https://doi.org/10.1136/bmj.328.7455.1561>
56. Ritchie J, Spencer L, O'Connor W. *Qualitative Research Practice, A Guide for Social Science Students and Researchers*. London: Sage; 2003.
57. Hyland G, Hay-Smith J, Treharne G. Women's experiences of doing long-term pelvic floor muscle exercises for the treatment of pelvic organ prolapse symptoms. *Int Urogynecol J* 2014;**25**:265–71. <https://doi.org/10.1007/s00192-013-2202-z>
58. Loohuis AMM, Wessels NJ, Jellema P, Vermeulen KM, Slieker-Ten Hove MC, van Gemert-Pijnen JEWC, et al. The impact of a mobile application-based treatment for urinary incontinence in adult women: design of a mixed-methods randomized controlled trial in a primary care setting. *Neurourol Urodyn* 2018;**37**:2167–76. <https://doi.org/10.1002/nau.23507>
59. Hespen van A, Heitner AP, Rock HM. *Incondition: Development and Process Evaluation of a Training Programme To Prevent or Reduce Urinary Incontinence in Older Women in Homes for the Elderly*. International Continence Society (ICS/UGA), 34th Annual Meeting, 23–27 August 2004, Paris, France, abstract number 427.
60. Yin RK. Case Study Research: Design and Methods. Bickman L, Rog DJ, editors. *Essential Guide to Qualitative Methods in Organizational Research*. London: Sage; 2009. p. 219.
61. Yin RK. *Case Study Research: Design and Methods*. 5th edn. Thousand Oaks, CA: Sage; 2014.

62. Alvesson M, Sköldbberg K. *Reflexive Methodology: New Vistas for Qualitative Research*. 2nd edn. London: Sage; 2000.
63. Grant A, Treweek S, Dreischulte T, Foy R, Guthrie B. Process evaluations for cluster-randomised trials of complex interventions: a proposed framework for design and reporting. *Trials* 2013;**14**:15. <https://doi.org/10.1186/1745-6215-14-15>
64. Kovandžić M, Chew-Graham C, Reeve J, Edwards S, Peters S, Edge D, *et al.* Access to primary mental health care for hard-to-reach groups: from 'silent suffering' to 'making it work'. *Soc Sci Med* 2011;**72**:763–72. <https://doi.org/10.1016/j.socscimed.2010.11.027>
65. Stake RE. Case Studies. In Denzin NK, Lincoln YS, editors. *Handbook of Qualitative Research*. Thousand Oaks, CA: Sage; 1994. pp. 236–48.
66. Ritchie J, Lewis J, McNaughton Nicolls C, Ormston R. *Qualitative Research Practice*. 2nd edn. Thousand Oaks, CA: Sage; 2014.
67. Wells M, Williams B, Treweek S, Coyle J, Taylor J. Intervention description is not enough: evidence from an in-depth multiple case study on the untold role and impact of context in randomised controlled trials of seven complex interventions. *Trials*. 2012;**13**:95 <https://doi.org/10.1186/1745-6215-13-95>
68. Hay-Smith EJC, Dean SG. Overcoming inertia: narratives from women embarking on conservative management of urinary incontinence. *Aust New Zeal Cont J* 2012;**18**:121.
69. Dumoulin C, Hay-Smith J, Frawley H, McClurg D, Alewijnse D, Bo K, *et al.* 2014 consensus statement on improving pelvic floor muscle training adherence: International Continence Society 2011 State-of-the-Science Seminar. *Neurourol Urodyn* 2015;**34**:600–5. <https://doi.org/10.1002/nau.22796>
70. Hay-Smith EJ, Herderschee R, Dumoulin C, Herbison GP. Comparisons of approaches to pelvic floor muscle training for urinary incontinence in women. *Cochrane Database Syst Rev* 2011;**12**:CD009508. <https://doi.org/10.1002/14651858.CD009508>
71. McClurg D, Frawley H, Hay-Smith J, Dean S, Chen SY, Chiarelli P, *et al.* Scoping review of adherence promotion theories in pelvic floor muscle training – 2011 ICS state-of-the-science seminar research paper i of iv. *Neurourol Urodyn* 2015;**34**:606–14. <https://doi.org/10.1002/nau.22769>
72. Bradway C. Women's narratives of long-term urinary incontinence. *Urol Nurs* 2005;**25**:337–44.
73. Delarmelindo Rde C, Parada CM, Rodrigues RA, Bocchi SC. Between suffering and hope: rehabilitation from urinary incontinence as an intervening component. *Cien Saude Colet* 2013;**18**:1981–91. <https://doi.org/10.1590/S1413-81232013000700013>
74. Hamid TA, Pakgothar M, Ibrahim R, Dastjerdi MV. 'Stain in life': The meaning of urinary incontinence in the context of Muslim postmenopausal women through hermeneutic phenomenology. *Arch Gerontol Geriatr* 2015;**60**:514–21. <https://doi.org/10.1016/j.archger.2015.01.003>
75. Bate P. *Perspectives on Context*. 2014. URL: https://pdfs.semanticscholar.org/39a5/4ed4d7e2e6a90ee24db6522c475d7a05af7d.pdf?_ga=2.106199168.2079529133.1544714634-2078294441.1544714634 (accessed 6 December 2018).
76. Pfadenhauer L, Rohwer A, Burns J, Booth A, Lysdahl KB, Hofmann B, *et al.* *Guidance for the Assessment of Context and Implementation in Health Technology Assessments (HTA) and Systematic Reviews of Complex Interventions: The Context and Implementation of Complex Interventions (CICI) Framework*. URL: www.integrate-hta.eu/wp-content/uploads/2016/02/Guidance-for-the-Assessment-of-Context-and-Implementation-in-HTA-and-Systematic-Reviews-of-Complex-Interventions-The-Co.pdf (accessed 6 December 2018).

77. Hay-Smith EJC, Ryan K, Dean S. The silent, private exercise: experiences of pelvic floor muscle training in a sample of women with stress urinary incontinence. *Physiotherapy* 2007;**93**:53–61. <https://doi.org/10.1016/j.physio.2006.10.005>
78. O’Cathain A, Murphy E, Nicholl J. Three techniques for integrating data in mixed methods studies. *BMJ* 2010;**341**:c4587. <https://doi.org/10.1136/bmj.c4587>
79. Williams KS, Assassa RP, Gillies CL, Abrams KR, Turner DA, Shaw C, *et al.* A randomized controlled trial of the effectiveness of pelvic floor therapies for urodynamic stress and mixed incontinence. *BJU Int* 2006;**98**:1043–50. <https://doi.org/10.1111/j.1464-410X.2006.06484.x>
80. Ibrahim IK, Hameed MMA, Taher EM, Shaheen EM, Elsayy MSAG. Efficacy of biofeedback-assisted pelvic floor muscle training in females with pelvic floor dysfunction. *Alexandria J Med* 2015;**51**:137–42. <https://doi.org/10.1016/j.ajme.2014.06.001>
81. Ahadi T, Taghvadoost N, Aminimoghaddam S, Forogh B, Bazazbehbahani R, Raissi GR. Efficacy of biofeedback on quality of life in stages I and II pelvic organ prolapse: a pilot study. *Eur J Obstet Gynecol Reprod Biol* 2017;**215**:241–6. <https://doi.org/10.1016/j.ejogrb.2017.06.023>
82. Liu L, Zhang Y, Gong J, Chen X, Wu H, Zhu W. Effects of different treatment methods on the clinical and urodynamic state of perimenopausal women with stress urinary incontinence. *Iran J Public Health* 2018;**47**:1090–7.
83. Terlikowski R, Dobrzycka B, Kinalski M, Kuryliszyn-Moskal A, Terlikowski SJ. Transvaginal electrical stimulation with surface-EMG biofeedback in managing stress urinary incontinence in women of premenopausal age: a double-blind, placebo-controlled, randomized clinical trial. *Int Urogynecol J* 2013;**24**:1631–8. <https://doi.org/10.1007/s00192-013-2071-5>
84. Hirakawa T, Suzuki S, Kato K, Gotoh M, Yoshikawa Y. Randomized controlled trial of pelvic floor muscle training with or without biofeedback for urinary incontinence. *Int Urogynecol J* 2013;**24**:1347–54. <https://doi.org/10.1007/s00192-012-2012-8>
85. Fitz FF, Resende APM, Stüpp L, Costa TF, Sartori MGF, Girão MJBC, *et al.* Efeito da adição do biofeedback ao treinamento dos músculos do assoalho pélvico para tratamento da incontinência urinária de esforço. *Rev Bras Ginecol e Obs* 2012;**34**:505–10. <https://doi.org/10.1590/S0100-72032012001100005>
86. Fitz FF, Stüpp L, da Costa TF, Bortolini MAT, Girão MJBC, Castro RA. Outpatient biofeedback in addition to home pelvic floor muscle training for stress urinary incontinence: a randomized controlled trial. *Neurourol Urodyn* 2017;**36**:2034–43. <https://doi.org/10.1002/nau.23226>
87. Bertotto A, Schwartzman R, Uchôa S, Wender MCO. Effect of electromyographic biofeedback as an add-on to pelvic floor muscle exercises on neuromuscular outcomes and quality of life in postmenopausal women with stress urinary incontinence: a randomized controlled trial. *Neurourol Urodyn* 2017;**36**:2142–7. <https://doi.org/10.1002/nau.23258>
88. Özlü A, Yıldız N, Öztekin Ö. Comparison of the efficacy of perineal and intravaginal biofeedback assisted pelvic floor muscle exercises in women with urodynamic stress urinary incontinence. *Neurourol Urodyn* 2017;**36**:2132–41. <https://doi.org/10.1002/nau.23257>
89. Abrams P, Cardozo L, Wagg A, Wein A, editors. *Incontinence*. 6th edn. Bristol: International Incontinence Society; 2017.
90. Dumoulin C, Cacciari LP, Hay-Smith EJC. Pelvic floor muscle training versus no treatment, or inactive control treatments, for urinary incontinence in women [published online ahead of print 10 04 2018]. *Cochrane Database Syst Rev* 2018. <https://doi.org/10.1002/14651858.CD005654.pub4>
91. Sran M, Mercier J, Wilson P, Lieblisch P, Dumoulin C. Physical therapy for urinary incontinence in postmenopausal women with osteoporosis or low bone density: a randomized controlled trial. *Menopause* 2016;**23**:286–93. <https://doi.org/10.1097/GME.0000000000000594>

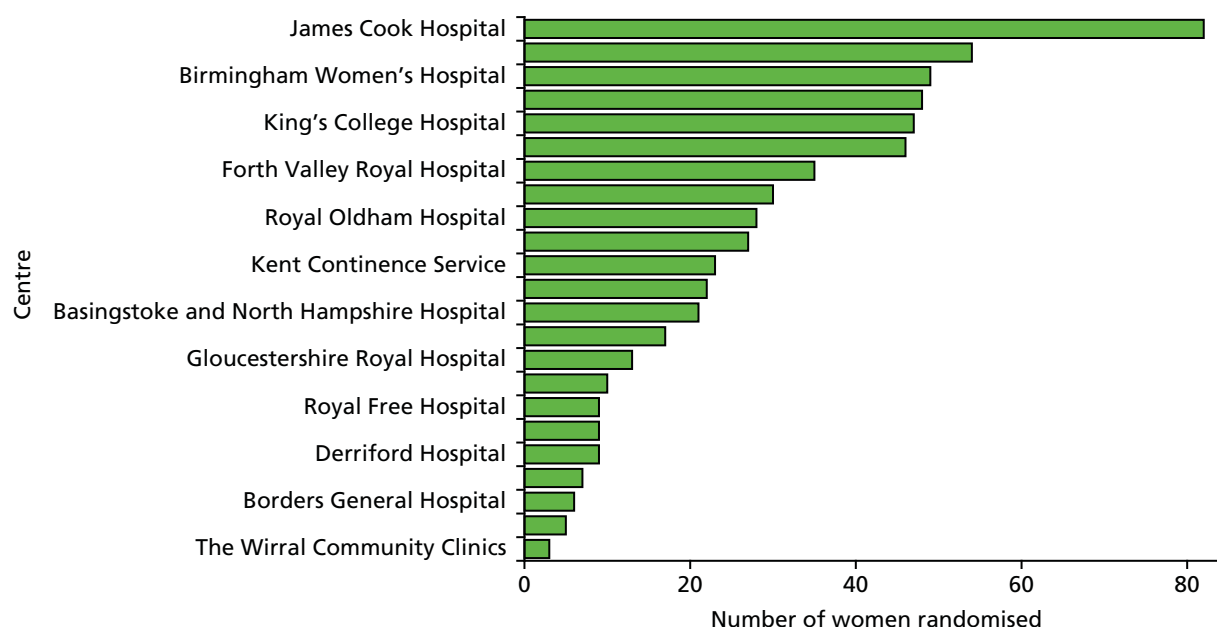
Appendix 1 Trial components



Appendix 2 Overview of participant data collected via questionnaires

Data collected	Time point			
	Baseline	6 months	12 months	24 months
Urinary outcomes				
• ICIQ-UI SF	X	X	X	X
• PGI-S	X			
• PGI-I		X	X	X
• Uptake of surgery for UI		X	X	X
• Uptake of other treatment for UI		X	X	X
• Other urinary symptoms (ICIQ-FLUTS)	X	X	X	X
Quality-of-life outcomes				
• UI-specific quality of life (ICIQ-FLUTSqol)	X	X	X	X
• General health (EQ-5D-3L)	X	X	X	X
Pelvic floor-related outcomes				
• Prolapse symptoms (POP-SS)	X	X	X	X
• Bowel symptoms (early version of International Consultation on Incontinence Questionnaire Bowel Short Form)	X	X	X	X
• Self-efficacy for PFMT	X	X	X	X
• Adherence to PFMT	X	X	X	X
Economic outcomes				
• Cost and use of NHS services			X	X
• Cost to the women and their families/carers				X

Appendix 3 Recruitment by centre



Appendix 4 Reasons for ineligibility

There were 51 women who were screened who were ineligible.

Ineligible

Unable to contract their pelvic floor muscles ($n = 18$).

Known nickel allergy or sensitivity ($n = 16$).

Prolapse stage 2, pelvic cancer, cognitive impairment or neurological disease ($n = 7$).

Urgency UI alone ($n = 6$).

Previous formal instruction in PFMT ($n = 3$).

Pregnant or < 1 year postnatal ($n = 1$).

Antimuscarinic medication ($n = 0$).

Currently participating in other research relating to their UI ($n = 0$).

Appendix 5 Sensitivity analysis under assumptions of missing not at random

A pattern mixture modelling approach was used to explore the effect of missing data under MNAR assumptions. Missing ICIQ-UI SF values were assumed to be between 2.5 units lower and 2.5 units greater, on average, than the observed values in (1) the biofeedback PFMT group only, (2) the basic PFMT group only and (3) in both groups. It was considered that 2.5 points was a minimum clinically important difference and, hence, a meaningful systematic difference to test in the sensitivity analyses. The pattern mixture model (Figure 12) tested data, assuming values of delta equal to -2.5 , -1.5 , -0.5 , $+0.5$, $+1.5$ and $+2.5$. None of the analyses produced a statistically significant treatment effect and therefore this sensitivity analysis did not provide any finding contradictory to the primary analysis.

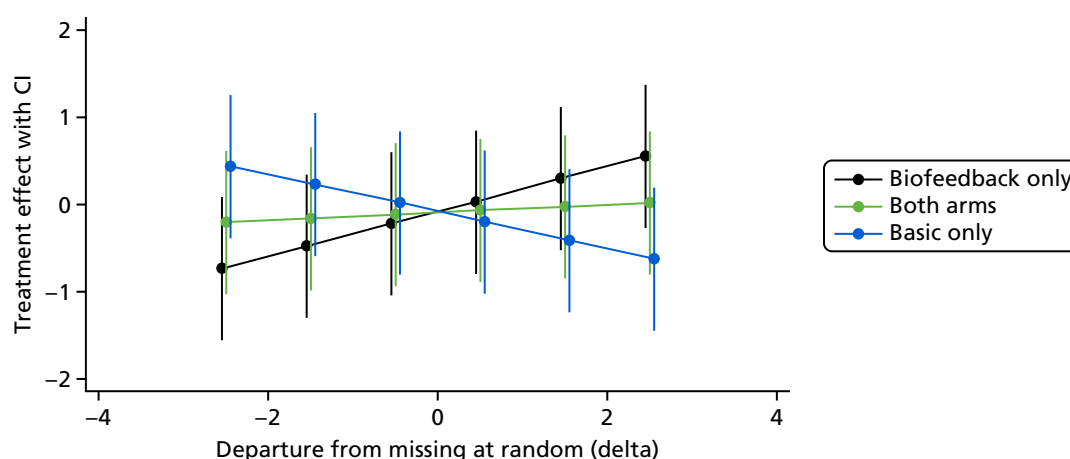


FIGURE 12 Pattern mixture model under assumptions of MNAR. Reproduced with permission from Hagen *et al.*³⁸ This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>. The figure includes minor additions and formatting changes to the original figure.

Appendix 6 Subgroup analysis

Variable	Subgroup	<i>n</i>	Mean	SD	Effect estimate	95% CI	<i>p</i> -value
Overall ICIQ–UI SF	Basic PFMT	235	8.5	4.9	−0.09	−0.92 to 0.75	0.84
	Biofeedback PFMT	225	8.2	5.1			
Type of incontinence	SUI	182	7.4	4.8	−0.80	−2.13 to 0.53	0.24
	MUI	278	8.9	5.0	0.38	−0.69 to 1.45	0.49
	Interaction				1.18	−0.53 to 2.89	0.18
Age (years)	< 50	265	8.3	5.1	−0.23	−1.32 to 0.87	0.68
	≥ 50	195	8.3	4.8	0.12	−1.17 to 1.40	0.86
	Interaction				0.34	−1.35 to 2.03	0.69
Severity of incontinence ^a	Mild/moderate	206	6.6	3.8	−0.75	−2.03 to 0.53	0.25
	Severe	254	9.7	5.4	0.28	−0.87 to 1.43	0.63
	Interaction				1.03	−0.69 to 2.75	0.24
Type of therapist	Physiotherapist	406	8.1	4.8	−0.24	−1.12 to 0.64	0.60
	Not physiotherapist	54	9.9	5.7	1.15	−1.38 to 3.68	0.37
	Interaction				−1.39	−4.07 to 1.29	0.31

a An ICIQ–UI SF score of < 13 is mild/moderate; an ICIQ–UI SF score of ≥ 13 is severe.

Appendix 7 Patient Global Impression of Improvement in urinary incontinence

Time point	Response	Treatment group					
		Biofeedback PFMT			Basic PFMT		
		N	n	%	N	n	%
6 months	Very much better	219	37	16.9	221	37	16.7
	Much better	219	59	26.9	221	48	21.7
	A little better	219	73	33.3	221	83	37.6
	No change	219	34	15.5	221	46	20.8
	A little worse	219	12	5.5	221	7	3.2
	Much worse	219	4	1.8	221	0	0
	Very much worse	219	0	0	221	0	0
12 months	Very much better	249	39	15.7	250	36	14.4
	Much better	249	62	24.9	250	56	22.4
	A little better	249	79	31.7	250	81	32.4
	No change	249	51	20.5	250	52	20.8
	A little worse	249	13	5.2	250	17	6.8
	Much worse	249	4	1.6	250	7	2.8
	Very much worse	249	1	0.4	250	1	0.4
24 months	Very much better	227	43	18.9	236	42	17.8
	Much better	227	50	22	236	48	20.3
	A little better	227	58	25.6	236	54	22.9
	No change	227	41	18.1	236	47	19.9
	A little worse	227	22	9.7	236	32	13.6
	Much worse	227	8	3.5	236	10	4.2
	Very much worse	227	5	2.2	236	3	1.3

Appendix 8 Bowel and prolapse symptoms

TABLE 35 Bowel symptoms

Symptom response	Treatment group					
	Biofeedback PFMT			Basic PFMT		
	<i>N</i>	<i>n</i>	%	<i>N</i>	<i>n</i>	%
Baseline						
<i>Difficulty emptying bowels?</i>						
Never	289	85	29.4	296	79	26.7
Occasionally	289	101	34.9	296	94	31.8
Sometimes	289	68	23.5	296	83	28
Most of the time	289	25	8.7	296	26	8.8
All of the time	289	10	3.5	296	14	4.7
<i>Rush to toilet?</i>						
Never	289	104	36	296	111	37.5
Occasionally	289	97	33.6	296	102	34.5
Sometimes	289	62	21.5	296	55	18.6
Most of the time	289	18	6.2	296	22	7.4
All of the time	289	8	2.8	296	6	2
<i>Stool leak?</i>						
Never	289	213	73.7	295	217	73.6
Occasionally	289	52	18	295	45	15.3
Sometimes	289	19	6.6	295	30	10.2
Most of the time	289	4	1.4	295	3	1
All of the time	289	1	0.3	295	0	0
<i>How often bowels open?</i>						
Three or more times a day	286	24	8.4	297	29	9.8
Twice a day	286	60	21	297	61	20.5
Once a day	286	139	48.6	297	125	42.1
Two or three times a week	286	53	18.5	297	74	24.9
Once a week or less	286	10	3.5	297	8	2.7
<i>Motions usually?</i>						
Watery	284	3	1.1	293	3	1
Sloppy	284	29	10.2	293	32	10.9
Soft and formed	284	216	76.1	293	198	67.6
Hard	284	36	12.7	293	60	20.5

continued

TABLE 35 Bowel symptoms (continued)

Symptom response	Treatment group					
	Biofeedback PFMT			Basic PFMT		
	<i>N</i>	<i>n</i>	%	<i>N</i>	<i>n</i>	%
6 months						
<i>Difficulty emptying bowels?</i>						
Never	180	66	36.7	176	49	27.8
Occasionally	180	59	32.8	176	51	29
Sometimes	180	33	18.3	176	54	30.7
Most of the time	180	14	7.8	176	14	8
All of the time	180	8	4.4	176	8	4.5
<i>Rush to toilet?</i>						
Never	180	76	42.2	176	70	39.8
Occasionally	180	68	37.8	176	67	38.1
Sometimes	180	28	15.6	176	31	17.6
Most of the time	180	5	2.8	176	6	3.4
All of the time	180	3	1.7	176	2	1.1
<i>Stool leak?</i>						
Never	180	136	75.6	176	137	77.8
Occasionally	180	33	18.3	176	27	15.3
Sometimes	180	10	5.6	176	11	6.3
All of the time	180	0	0	176	1	0.6
Most of the time	180	1	0.6	176	0	0
<i>How often bowels open?</i>						
Three or more times a day	178	15	8.4	176	13	7.4
Twice a day	178	29	16.3	176	37	21
Once a day	178	98	55.1	176	85	48.3
Two or three times a week	178	29	16.3	176	35	19.9
Once a week or less	178	7	3.9	176	6	3.4
<i>Motions usually?</i>						
Watery	178	1	0.6	173	2	1.2
Sloppy	178	14	7.9	173	12	6.9
Soft and formed	178	134	75.3	173	126	72.8
Hard	178	29	16.3	173	33	19.1
12 months						
<i>Difficulty emptying bowels?</i>						
Never	187	62	33.2	183	50	27.3
Occasionally	187	67	35.8	183	64	35
Sometimes	187	44	23.5	183	43	23.5
Most of the time	187	12	6.4	183	20	10.9

TABLE 35 Bowel symptoms (continued)

Symptom response	Treatment group					
	Biofeedback PFMT			Basic PFMT		
	<i>N</i>	<i>n</i>	%	<i>N</i>	<i>n</i>	%
All of the time	187	2	1.1	183	6	3.3
<i>Rush to toilet?</i>						
Never	188	80	42.6	182	83	45.6
Occasionally	188	75	39.9	182	65	35.7
Sometimes	188	27	14.4	182	26	14.3
Most of the time	188	4	2.1	182	6	3.3
All of the time	188	2	1.1	182	2	1.1
<i>Stool leak?</i>						
Never	188	144	76.6	182	145	79.7
Occasionally	188	35	18.6	182	26	14.3
Sometimes	188	6	3.2	182	9	4.9
Most of the time	188	2	1.1	182	1	0.5
All of the time	188	1	0.5	182	1	0.5
<i>How often bowels open?</i>						
Three or more times a day	187	11	5.9	181	13	7.2
Twice a day	187	39	20.9	181	38	21
Once a day	187	100	53.5	181	84	46.4
Two or three times a week	187	29	15.5	181	39	21.5
Once a week or less	187	8	4.3	181	7	3.9
<i>Motions usually?</i>						
Watery	185	3	1.6	178	1	0.6
Sloppy	185	11	5.9	178	11	6.2
Soft and formed	185	150	81.1	178	130	73
Hard	185	21	11.4	178	36	20.2
24 months						
<i>Difficulty emptying bowels?</i>						
Never	161	59	36.6	167	37	22.2
Occasionally	161	64	39.8	167	69	41.3
Sometimes	161	25	15.5	167	44	26.3
Most of the time	161	10	6.2	167	12	7.2
All of the time	161	3	1.9	167	5	3
<i>Rush to toilet?</i>						
Never	162	71	43.8	168	70	41.7
Occasionally	162	54	33.3	168	60	35.7
Sometimes	162	30	18.5	168	26	15.5

continued

TABLE 35 Bowel symptoms (continued)

Symptom response	Treatment group					
	Biofeedback PFMT			Basic PFMT		
	<i>N</i>	<i>n</i>	%	<i>N</i>	<i>n</i>	%
Most of the time	162	4	2.5	168	10	6
All of the time	162	3	1.9	168	2	1.2
<i>Stool leak?</i>						
Never	162	129	79.6	167	125	74.9
Occasionally	162	24	14.8	167	29	17.4
Sometimes	162	8	4.9	167	9	5.4
Most of the time	162	1	0.6	167	4	2.4
<i>How often bowels open?</i>						
Three or more times a day	161	11	6.8	168	15	8.9
Twice a day	161	31	19.3	168	36	21.4
Once a day	161	82	50.9	168	68	40.5
Two or three times a week	161	28	17.4	168	41	24.4
Once a week or less	161	9	5.6	168	8	4.8
<i>Motions usually?</i>						
Watery	159	2	1.3	167	3	1.8
Sloppy	159	13	8.2	167	16	9.6
Soft and formed	159	130	81.8	167	118	70.7
Hard	159	14	8.8	167	30	18

Reproduced with permission from Hagen *et al.*³⁸ This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>. The table includes minor additions and formatting changes to the original table.

TABLE 36 Prolapse symptoms (POP-SS)

Time point	Treatment group					
	Biofeedback PFMT			Basic PFMT		
	<i>n</i>	Mean	SD	<i>n</i>	Mean	SD
Baseline	274	6.4	5.7	286	6.7	5.6
6 months	174	4.9	5.1	167	5.0	4.9
12 months	176	4.6	4.8	172	5.0	5.1
24 months	157	4.5	5.0	161	4.9	5.0

Reproduced with permission from Hagen *et al.*³⁸ This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>. The table includes minor additions and formatting changes to the original table.

Appendix 9 Estimation of unit cost of biofeedback machines per participant

Equipment	Year	Equipment type	Cost (£)	HCHS in 2017 (£)	Average total cost (£)
Biofeedback unit	2013	NeuroTrac	89.95	99.37	
	2014	NeuroTrac	104.95	113.89	
	2014	Peritone (Neen; Performance Health, Huthwaite, UK)	125.84	136.56	
	2014	Peritone	113.49	123.16	
	2015	Simplex Plus	94.00	100.30	114.66
Two packets of electrodes	2013	Reusable	5.00	5.52	
	2014	Reusable	3.75	4.07	
	2014	Reusable	4.25	4.61	
	2014	Reusable	9.90	10.74	
	2015	Reusable	13.50	14.41	7.87
One probe	2013	Periform® (Neen; Performance Health, Huthwaite, UK)	14.75	16.29	
	2014	Periform	13.50	14.65	
	2015	Periform	9.90	10.56	13.84
Laptop	2013	IBM thinkpad (IBM Corporation, Armonk, NY, USA)	390.00	430.84	
	2015	HP laptop (HP Inc., Palo Alto, CA, USA)	450.00	480.17	455.51
Software	2014	Licence	203.00	220.29	
	2014	Licence	182.31	197.84	
	2014	Licence	101.00	109.60	
	2014	Licence	122.67	133.12	165.22
Printer	2013	HP	45.00	49.71	
	2014	HP	75.00	81.39	65.55
					855.04
Printer cable	2013		3.00	3.31	
Eighteen spare reference leads for unit	2014		8.47	9.19	
Total cost per woman					42.92
Total cost per woman (discounted at 3.5%)					37.40

Appendix 10 Unit costs of patient-reported medications

Drug	Dose	Duration	Pack size (number of tablets)	Drug tariff (£)
SUI				
Duloxetine (Yentreve; Eli Lilly and Company, Indianapolis, IN, USA)	40 mg twice daily	Ongoing	56	5.45
Solifenacin (Vesicare)	5 mg initially, increased to 10 mg once daily	Ongoing	30	27.62 (5 mg); 35.91 (10 mg)
Mirabegron (Betmiga)	50 mg once daily	Ongoing	30	29.00
Oxybutynin hydrochloride (Lyrinel XL; Janssen Pharmaceutica, Beerse, Belgium)	5 mg (twice a day on average)	Ongoing	56	1.32
Tolterodine tartrate (Detrusitol; Pfizer Inc., New York, NY, USA)	2 mg twice daily (immediate release)	Ongoing	56	2.17
Tolterodine (Mariosea XL; Teva Pharmaceutical Industries Ltd, Petah Tikva, Israel) (all capsules)	4 mg once daily (moderate release)	Ongoing	28 (capsules)	11.60 (2-mg capsules); 25.78 (4-mg capsules)
Trospium chloride (Regurin; Contura Ltd, London, UK)	20 mg (immediate release) twice daily	Ongoing	60	8.34
	60 mg (modified release) once daily		28	
Fesoterodine fumarate (Toviaz; Pfizer Inc.)	4 mg once daily, increased if necessary up to 8 mg once daily	Ongoing	28	25.78
Darifenacin (Emselex; Merus Labs Luxco S.á R.L., Luxembourg)	7.5 mg once daily, increased to 15 mg after 2 weeks	Ongoing	28	25.48 (both 7.5 mg and 15 mg)
Urinary infections				
Trimethoprim (AAH Pharmaceuticals, Coventry, UK)	200 mg twice daily	Ongoing	14 (28)	0.73 (1.46)
Cefalexin (AAH Pharmaceuticals)	Antibacterial prophylaxis of recurrent urinary tract infection: 125 mg once a day	Ongoing	100	1.90
Antibiotics				
Nitrofurantoin (AAH Pharmaceuticals)	50 mg four times a day	5–7 days	28	11.36

Appendix 11 Unit cost of incontinence products (pads that women wear and disposable bed pads)

Pads	Pack size (n pads)	Cost per pack (£)	Cost per pad (£)
Interlude (Toiletry Sales Ltd, Wakefield, UK) incontinence pads	12	1.95	0.16
Interlude incontinence pads: extra	10	1.95	0.20
Interlude incontinence pads: extra plus	8	1.95	0.24
Suprem (Ontex Group, Aalst, Belgium) heavy incontinence pads	20	7.95	0.40
Suprem light incontinence pads	28	2.95	0.11
Suprem light incontinence pads: mini	20	1.45	0.07
Suprem moderate incontinence pads	25	5.95	0.24
Suprem severe incontinence pads	20	7.95	0.40
Suprem all in one incontinence pads: extra	24	10.95	0.46
Suprem all in one incontinence pads: maxi	20	10.95	0.55
Suprem all in one incontinence pads: super	22	11.65	0.53
Suprem all in one incontinence pads: extra large	20	12.95	0.65
Suprem all in one incontinence pads: extra large – maxi	20	11.25	0.56
Suprem all in one incontinence pads: large – extra	24	10.95	0.46
Suprem all in one incontinence pads: large – super	22	11.25	0.51
Suprem all in one incontinence pads: regular	26	10.95	0.42
Suprem all in one incontinence pads: small – extra	20	8.95	0.45
Suprem all in one incontinence pads: small – maxi	20	8.95	0.45
Average			0.38
Disposable bed pads (cm)			
60 × 90	25	7.95	0.32
60 × 60	25	5.95	0.24
40 × 60	25	4.25	0.14
60 × 75	25	6.45	0.26
Average			0.24

Appendix 12 Generalised linear model diagnostics test results

Family	Chi-squared	p-value
Gamma	0.1545	0.6942
Inverse Gaussian	1.2280	0.2678
Poisson	3.5887	0.0582
Gaussian	11.5304	0.0007
	Identity/gamma	Log/gamma
Pearson correlation test	0.5781	0.5991
Pregibon link test	0.1660	0.1436
Modified Hosmer–Lemeshow	0.2171	0.0486
QALYs	Chi-squared	p-value
Gaussian	10.5675	0.0012
Poisson	13.4584	0.0002
Gamma	16.6984	0.0000
Inverse Gaussian	20.2876	0.0000
	Identity/Gaussian	Log/Gaussian
Pearson correlation test	1.0000	0.9726
Pregibon link test	0.4330	0.3771
Modified Hosmer–Lemeshow	0.2751	0.2031

Appendix 13 Missing data (costs)

Variable	Treatment group, n/N (%)	
	Basic PFMT	Biofeedback PFMT
Treatment cost	2/300 (0.67)	5/300 (1.67)
GP visits		
6 months	117/300 (39.00)	117/300 (39.00)
12 months	114/300 (38.00)	113/300 (37.67)
24 months	168/300 (56.00)	175/300 (58.33)
Total GP visits	200/300 (66.67)	201/300 (67.00)
General practice nurse visits		
6 months	120/300 (40.00)	120/300 (40.00)
12 months	117/300 (39.00)	122/300 (40.67)
24 months	172/300 (57.33)	173/300 (57.67)
Total nurse visits	206/300 (68.67)	204/300 (68.00)
Hospital doctor visits		
6 months	137/300 (45.67)	144/300 (48.00)
12 months	119/300 (39.67)	115/300 (38.33)
24 months	169/300 (56.33)	176/300 (58.67)
Total hospital doctor visits	211/300 (70.33)	217/300 (72.33)
Nurse/physiotherapist visits		
6 months	121/300 (40.33)	123/300 (41.00)
12 months	119/300 (39.67)	119/300 (39.67)
24 months	171/300 (57.00)	176/300 (58.67)
Total nurse/physiotherapist visits	207/300 (69.00)	206/300 (68.67)
Hospital stay		
6 months	121/300 (40.33)	120/300 (40.00)
12 months	125/300 (41.67)	120/300 (40.00)
24 months	169/300 (56.33)	176/300 (58.67)
Total hospital stay	206/300 (68.67)	216/300 (72.00)
Surgical interventions		
6 months	136/300 (45.33)	128/300 (42.67)
12 months	90/300 (30.00)	96/300 (32.00)
24 months	138/300 (46.00)	146/300 (48.67)
Total surgical interventions	214/300 (71.33)	219/300 (73.00)
Medications		
6 months	96/300 (32.00)	91/300 (30.33)
12 months	68/300 (22.67)	66/300 (22.00)
24 months	126/300 (42.00)	132/300 (44.00)
Total medications	172/300 (57.33)	175/300 (58.33)
Total cost (missing either some intervention or follow-up data)	262/300 (87.33)	260/300 (86.67)

Appendix 14 Missing data (utility measures)

Variable	Treatment group, n/N (%)	
	Basic PFMT	Biofeedback PFMT
EQ-5D-3L scores		
Baseline	5/300 (1.67)	13/300 (4.33)
6 months	71/300 (23.67)	75/300 (25.00)
12 months	59/300 (19.67)	65/300 (21.67)
24 months	129/300 (43.00)	134/300 (44.67)
First-year QALY	98/300 (32.67)	108/300 (36.00)
Second-year QALY	158/300 (52.67)	174/300 (58.00)
ICIQ-LUTSqol scores		
Baseline	3/300 (1.00)	8/300 (2.67)
6 months	120/300 (40.00)	114/300 (38.00)
24 months	165/300 (55.00)	172/300 (57.33)
First-year QALY	147/300 (49.00)	140/300 (46.67)
Second-year QALY	200/300 (66.67)	202/300 (67.33)

Appendix 15 Proportion of participants experiencing any problems on each of the EQ-5D-3L domains: analysis based on all of the available EQ-5D-3L data points

Time point and group	Mobility, n/N (%)	Self-care, n/N (%)	Usual activities, n/N (%)	Pain/discomfort, n/N (%)	Anxiety/depression, n/N (%)
Baseline					
Basic PFMT	38/297 (12.79)	13/297 (4.38)	50/297 (16.84)	106/296 (35.81)	102/296 (34.46)
Biofeedback PFMT	48/289 (16.61)	19/289 (6.57)	57/289 (19.72)	114/288 (39.58)	105/288 (36.45)
6 months					
Basic PFMT	36/230 (15.65)	12/230 (6.52)	45/230 (19.57)	77/230 (33.48)	72/229 (31.44)
Biofeedback PFMT	39/227 (17.18)	21/228 (9.21)	50/227 (22.03)	76/228 (33.33)	94/228 (41.23)
12 months					
Basic PFMT	38/248 (15.32)	17/246 (6.91)	48/249 (19.28)	86/247 (34.82)	85/246 (34.55)
Biofeedback PFMT	48/241 (19.92)	32/244 (13.11)	56/244 (22.95)	95/242 (39.26)	93/241 (38.59)
24 months					
Basic PFMT	26/174 (14.94)	13/174 (7.47)	43/174 (24.71)	62/172 (36.05)	56/171 (32.75)
Biofeedback PFMT	38/168 (22.62)	18/168 (10.71)	37/168 (22.95)	50/168 (29.76)	56/167 (33.53)

Appendix 16 Proportion of participants who reported issues on each domain of the ICIQ-LUTSqol: analysis based on all of the available EQ-5D-3L data points

Time point and group	Role limitations, n/N (%)	Physical limitations, n/N (%)	Social limitations, n/N (%)	Emotional problems, n/N (%)	Sleep disturbance, n/N (%)
Baseline					
Basic PFMT	244/297 (82.15)	279/297 (93.94)	154/297 (51.85)	244/297 (82.15)	168/297 (56.57)
Biofeedback PFMT	246/292 (84.25)	282/292 (96.58)	165/292 (56.51)	244/292 (83.56)	182/292 (62.33)
6 months					
Basic PFMT	109/182 (59.89)	151/182(82.97)	62/180 (34.44)	112/180(62.22)	83/180 (46.11)
Biofeedback PFMT	120/186 (64.51)	161/186 (86.56)	67/186 (36.02)	119/186 (63.98)	77/186 (41.40)
12 months					
Basic PFMT	103/184 (55.98)	146/184 (79.34)	62/184(33.70)	115/184(62.50)	84/184 (45.65)
Biofeedback PFMT	109/189 (57.67)	158/189 (83.60)	68/189 (35.98)	114/189 (60.31)	86/189 (45.50)
24 months					
Basic PFMT	74/136 (54.41)	111/136 (81.62)	40/136 (29.41)	81/135 (60.00)	65/135 (48.15)
Biofeedback PFMT	69/129 (53.49)	108/130 (83.07)	43/132 (32.58)	78/132 (59.09)	51/132 (38.64)

Appendix 17 Sensitivity analysis cost-effectiveness acceptability curves and incremental cost-effectiveness ratio scatterplots

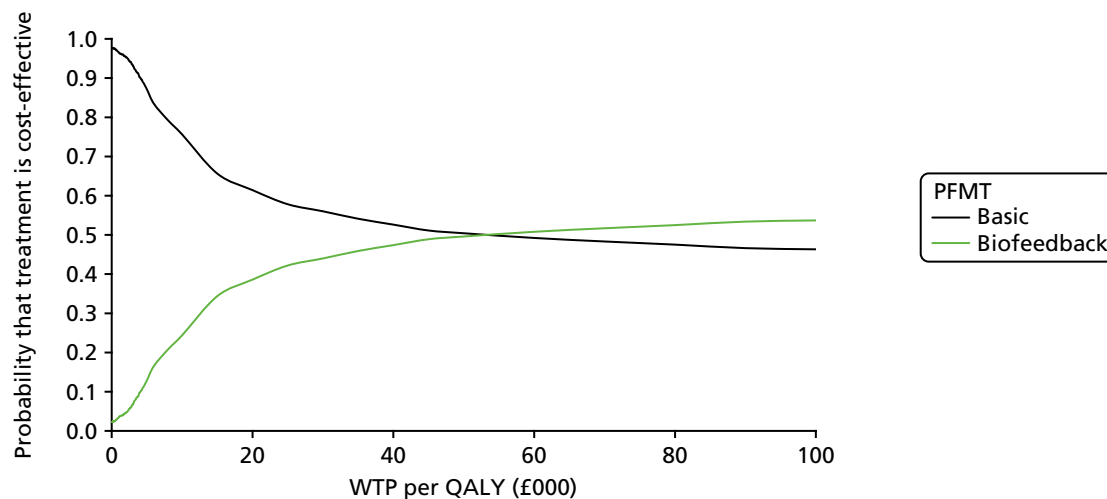


FIGURE 13 Cost-effectiveness acceptability curve: SUI subgroup analysis.

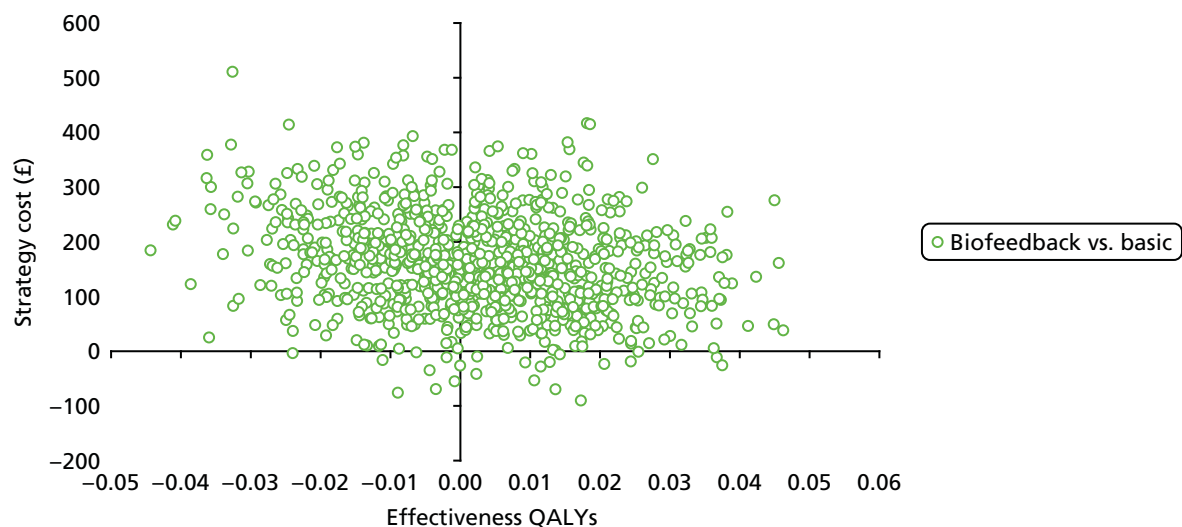


FIGURE 14 Scatterplot of incremental cost and incremental QALYs: SUI subgroup analysis.

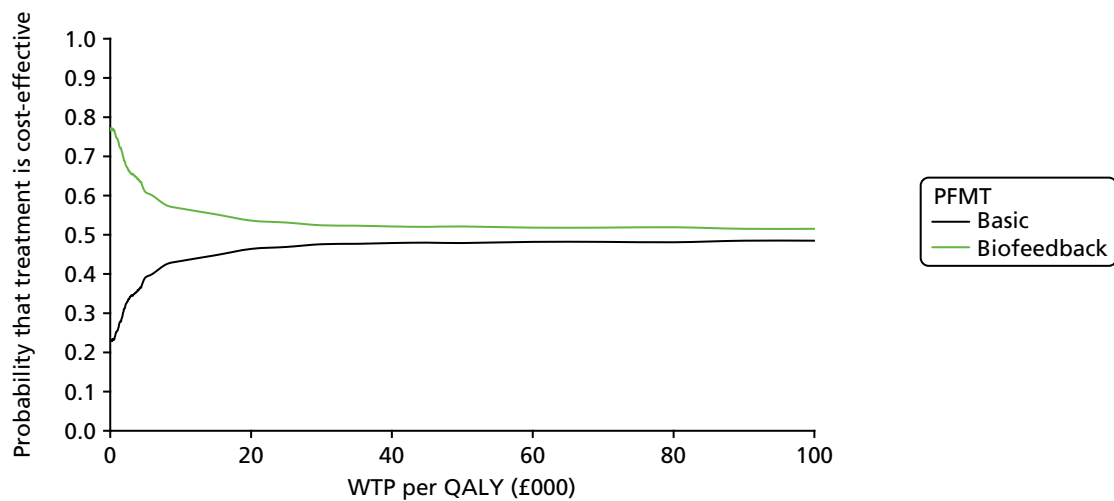


FIGURE 15 Cost-effectiveness acceptability curve: MUI subgroup analysis.

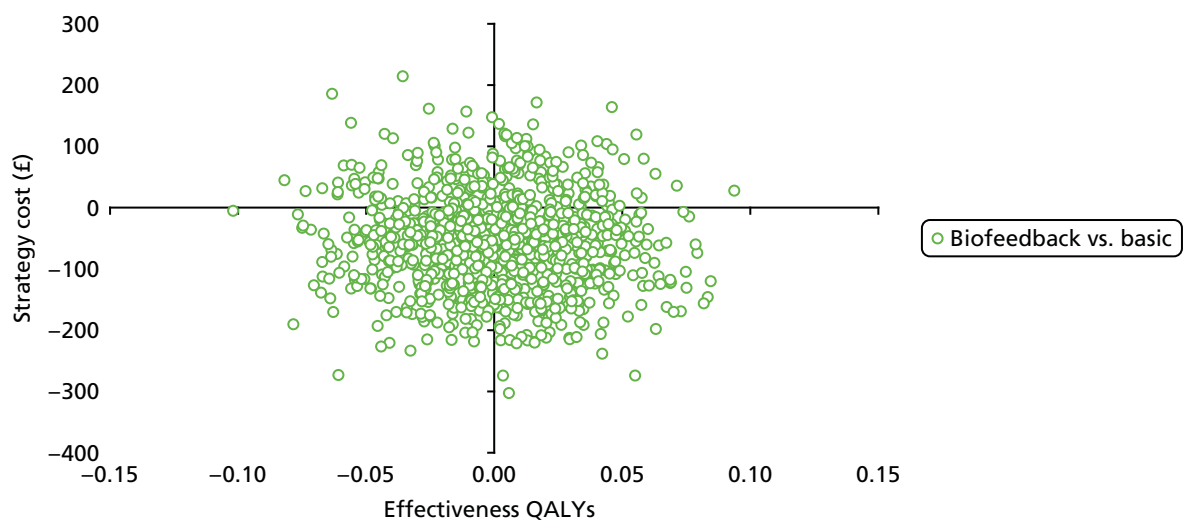


FIGURE 16 Scatterplot of incremental costs and incremental QALYs: MUI subgroup analysis.

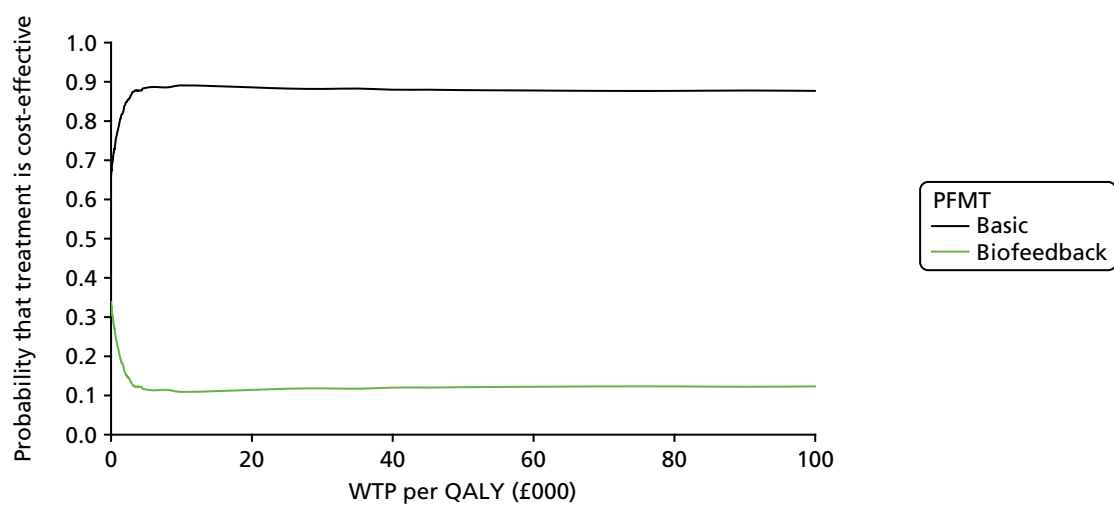


FIGURE 17 Cost-effectiveness acceptability curve: age < 50 years subgroup analysis.

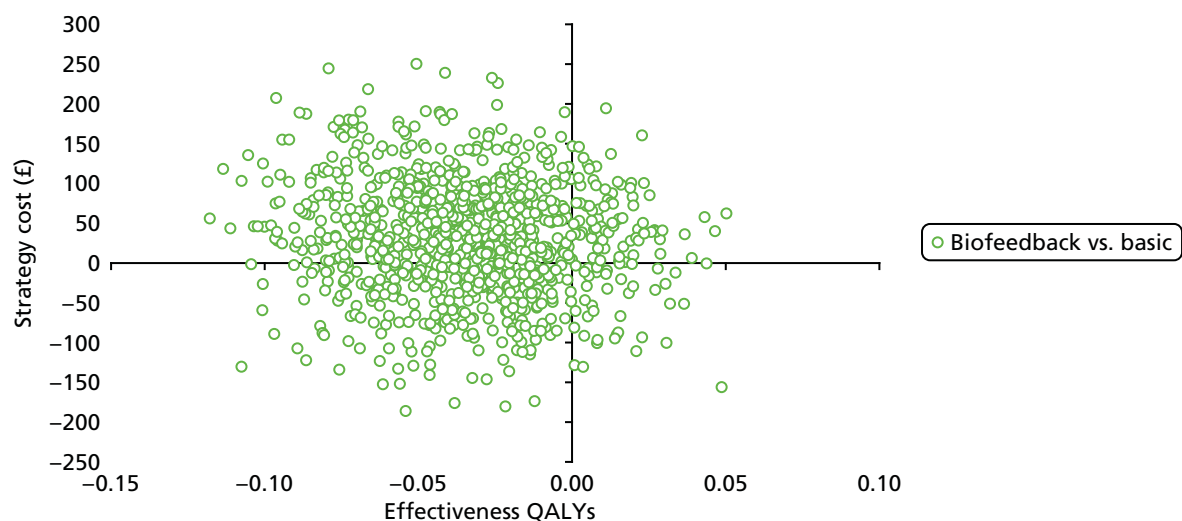


FIGURE 18 Scatterplot of incremental costs and QALYs: age < 50 years subgroup analysis.

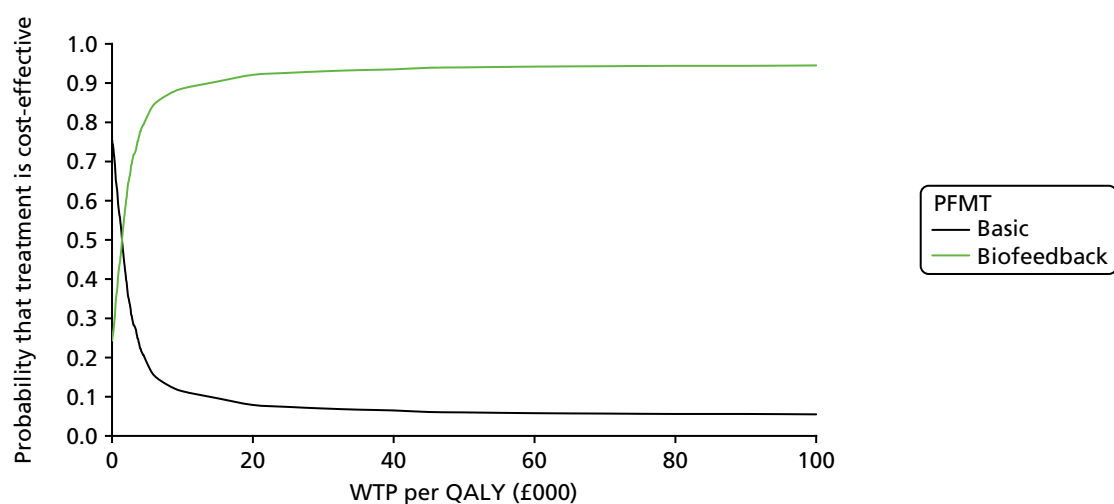


FIGURE 19 Cost-effectiveness acceptability curve: age ≥ 50 years subgroup analysis.

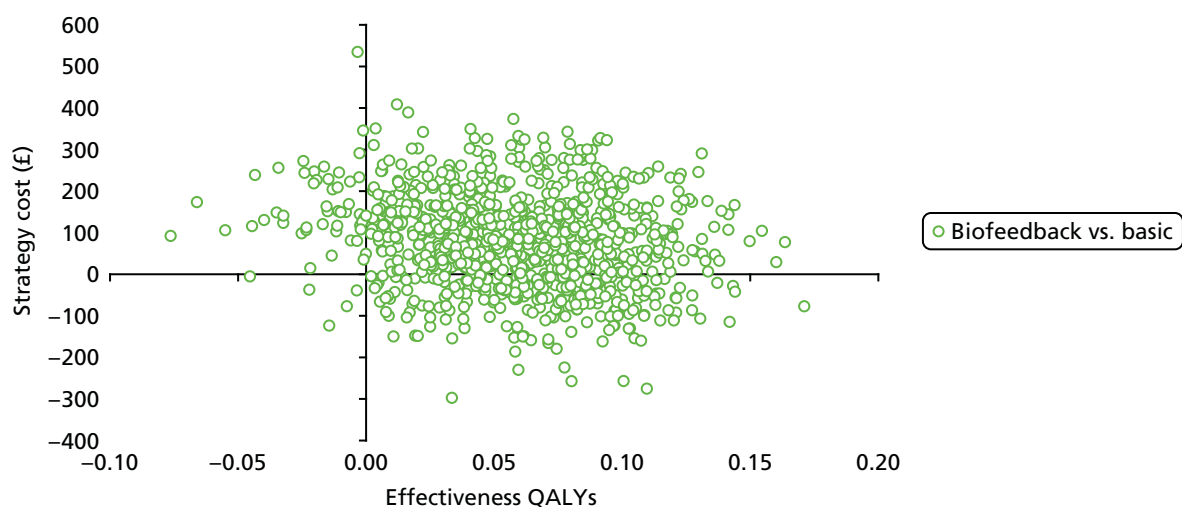


FIGURE 20 Scatterplot of incremental costs and incremental QALYs: age ≥ 50 years subgroup analysis.

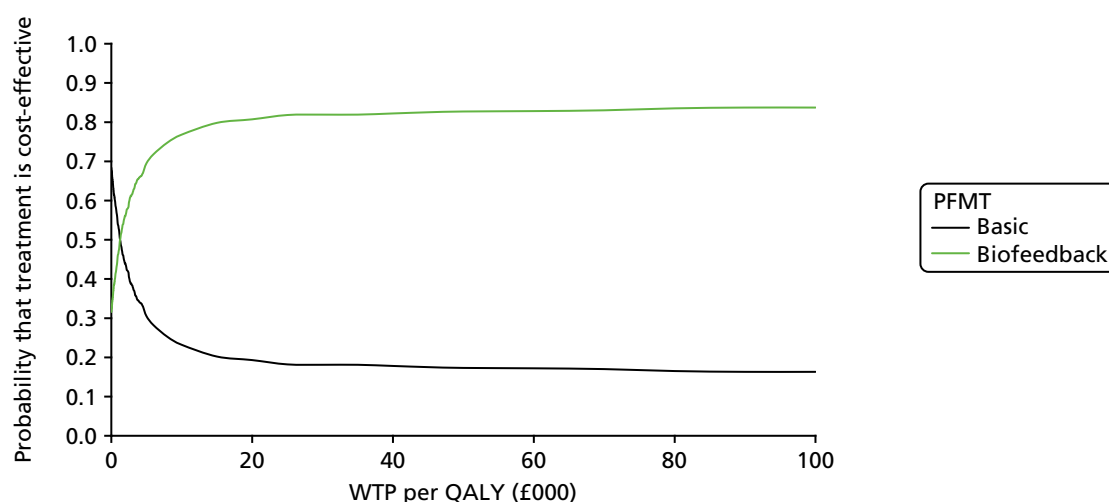


FIGURE 21 Cost-effectiveness acceptability curve: lower severity incontinence level (ICIQ-UI SF score of < 13) subgroup analysis.

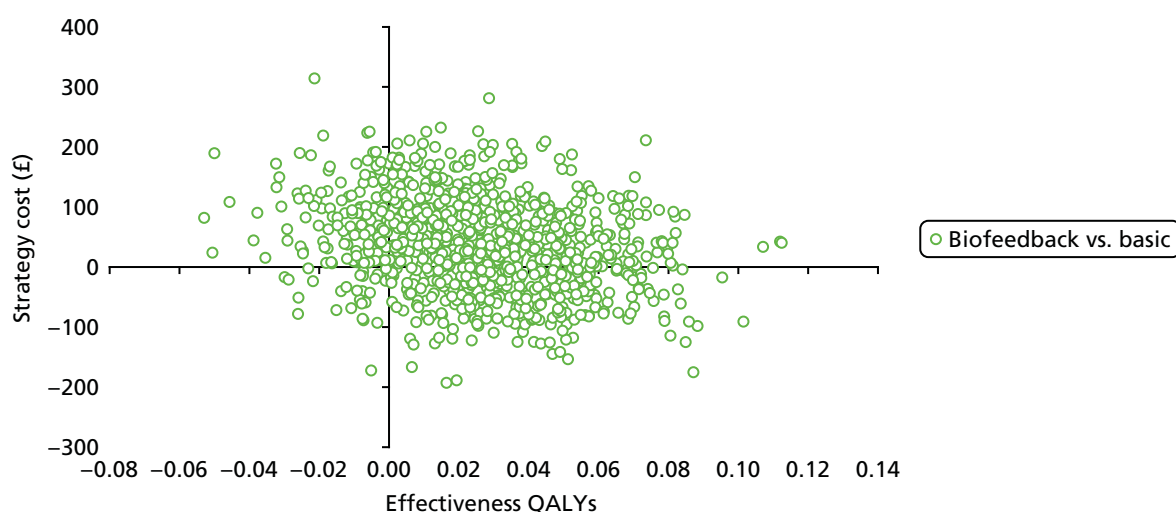


FIGURE 22 Scatterplot of incremental cost and incremental QALYs: lower severity incontinence level (ICIQ-UI SF score of < 13) subgroup analysis.

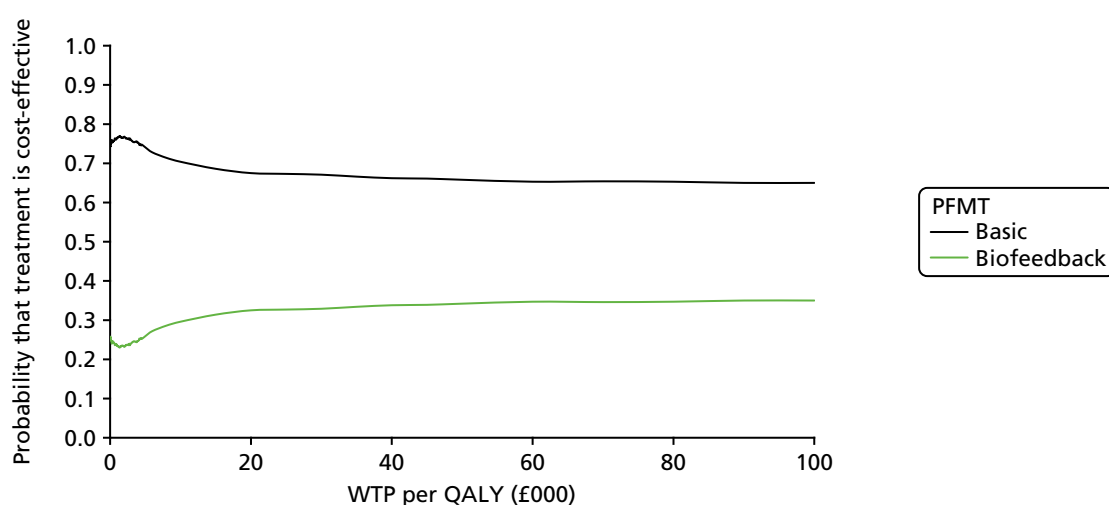


FIGURE 23 Cost-effectiveness acceptability curve: higher severity incontinence level (ICIQ-UI SF score of ≥ 13) subgroup analysis.

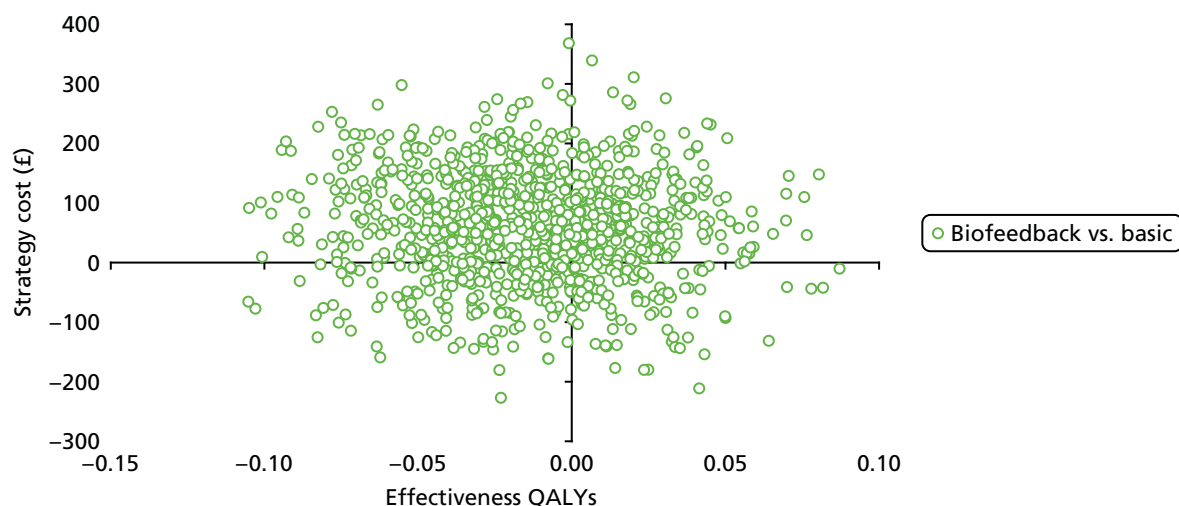


FIGURE 24 Scatterplot of incremental costs and incremental QALYs: higher severity incontinence level (ICIQ-UI SF score of ≥ 13) subgroup analysis.

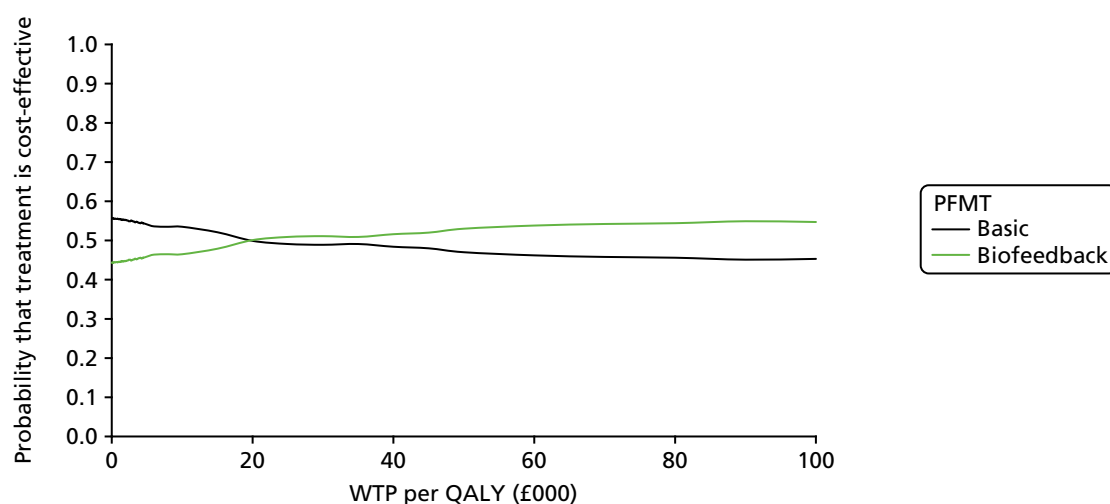


FIGURE 25 Cost-effectiveness acceptability curve: using ICIQ-LUTSqol scores as a utility measure.

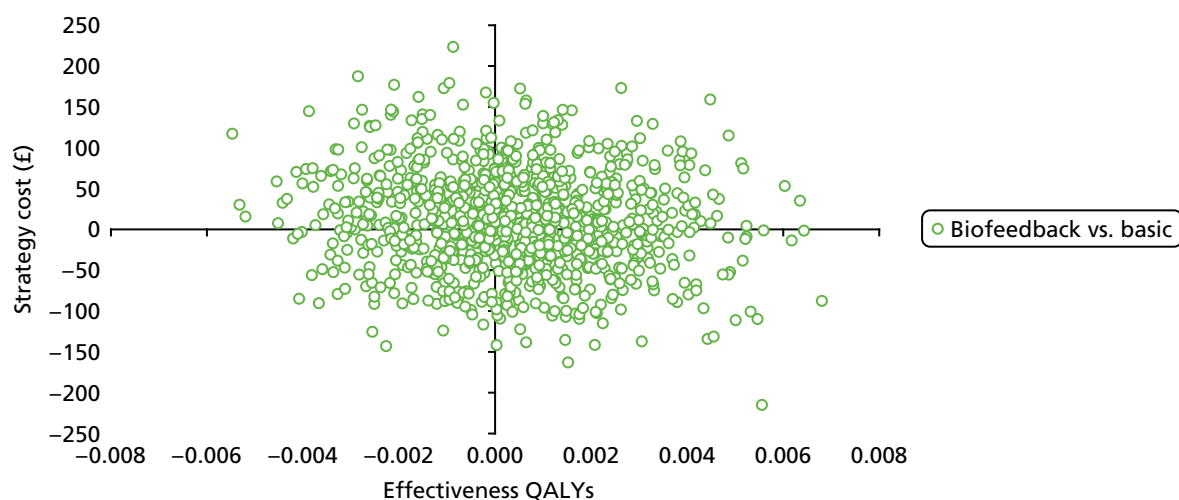


FIGURE 26 Scatterplot of incremental costs and incremental QALYs: using ICIQ-LUTSqol scores as a utility measure.

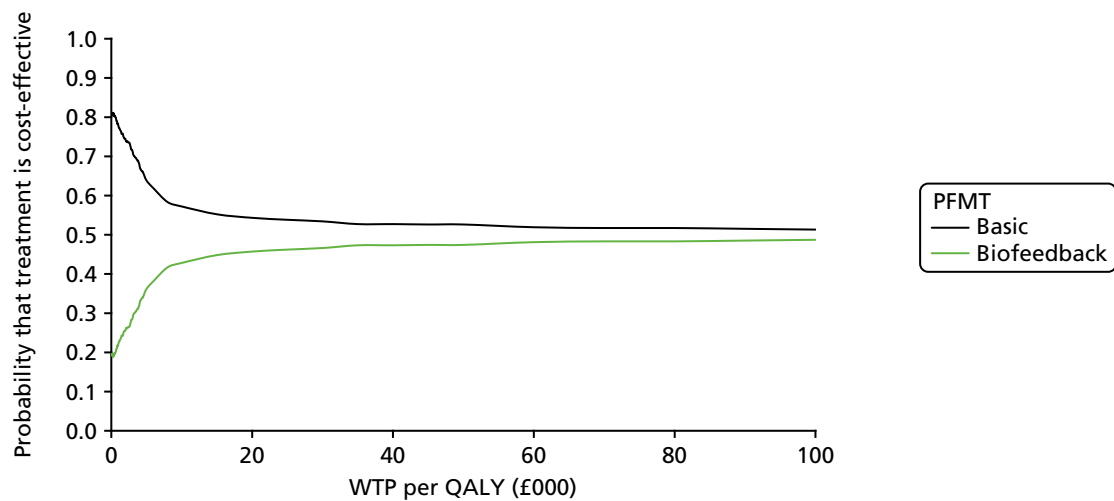


FIGURE 27 Cost-effectiveness acceptability curve: undiscounted costs and effects.

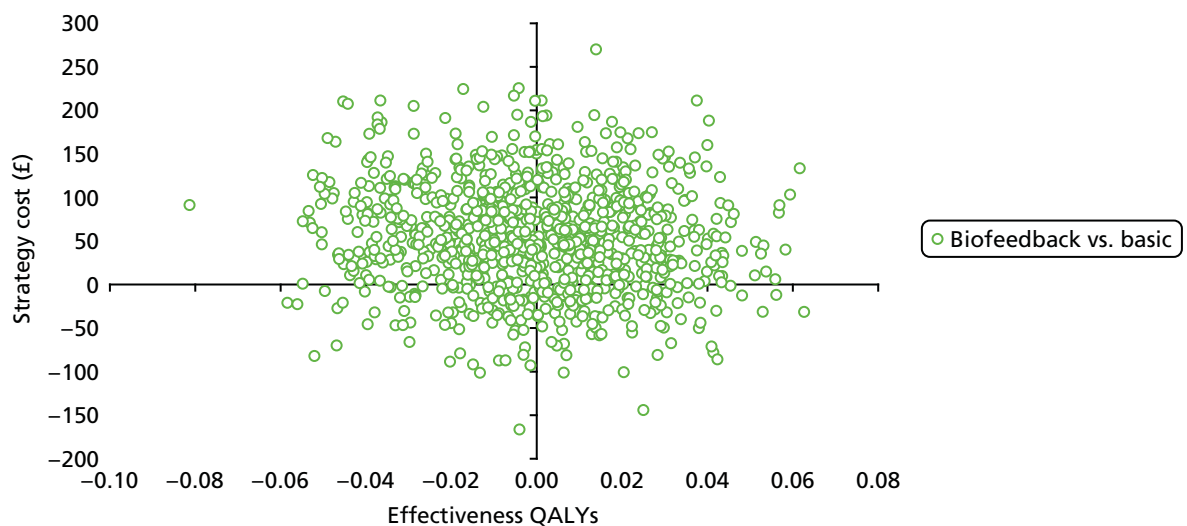


FIGURE 28 Scatterplot of incremental costs and incremental QALYs: undiscounted costs and effects.

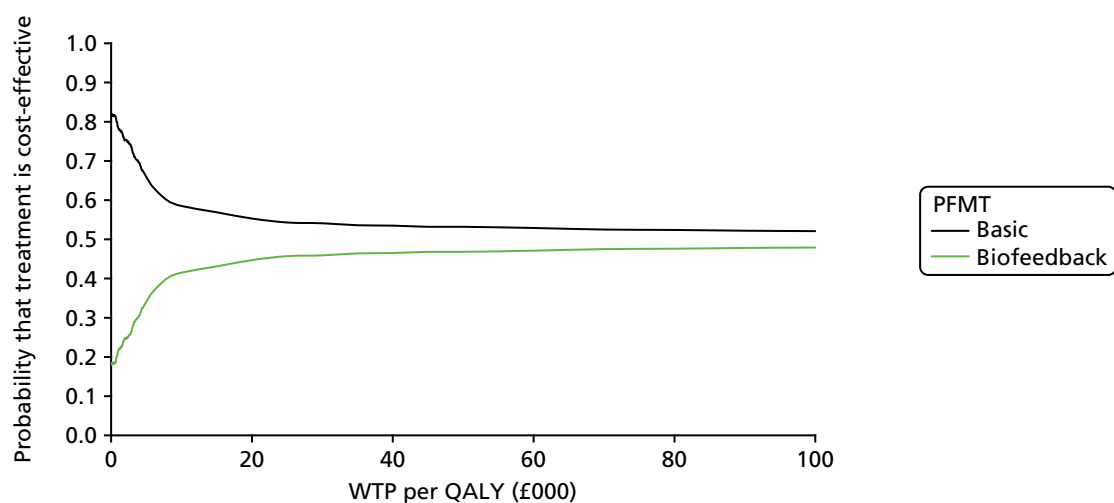


FIGURE 29 Cost-effectiveness acceptability curve: discounted costs and effects at 6%.

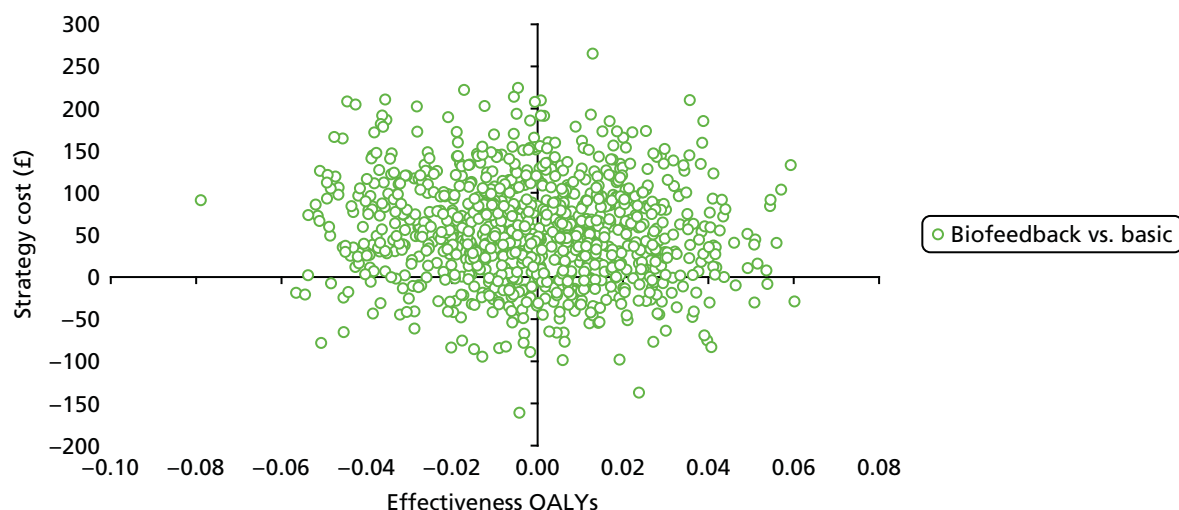


FIGURE 30 Scatterplot of incremental costs and incremental QALYs: discounted costs and effects at 6%.

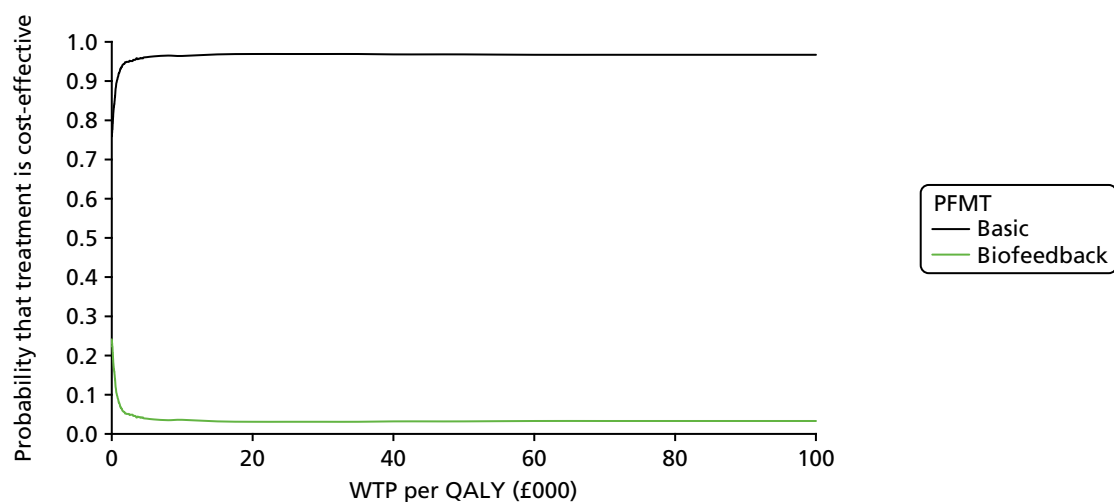


FIGURE 31 Cost-effectiveness acceptability curve: unadjusted cost-effectiveness analysis.

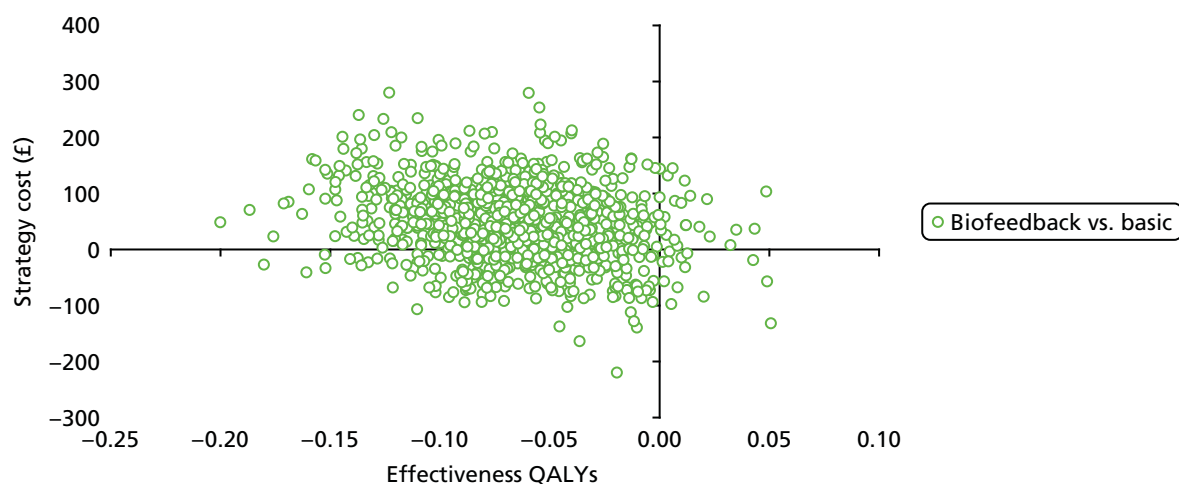


FIGURE 32 Scatterplot of incremental cost and incremental QALYs: unadjusted cost-effectiveness analysis.

Appendix 18 Examples of core and optional components of the basic PFMT and biofeedback PFMT intervention protocols, as worded on the therapist assessment form checklists

Appointment	Basic PFMT core component	Biofeedback core component
1	Practise 'the Knack' Record and both initial PFMT goal in exercise diary	Teach probe and electrode insertion/removal, turn biofeedback unit on/off Record and both initial biofeedback behaviour goal in exercise diary
6	Collect exercise diary Praise any PFMT achievements	Download biofeedback unit data and save Praise any biofeedback achievements
Optional component		
1	Suggest one fast contraction every time PFMT remembered Praise for intention to do home PFMT	Allay anxiety about biofeedback and its use Praise for intention to use biofeedback
6	Remark on disparity between PFMT goals and actions Offer information/teaching for self-feedback	Remark on disparity between biofeedback goal and actions Offer information/teaching of alternatives to biofeedback

EME
HS&DR
HTA
PGfAR
PHR

Part of the NIHR Journals Library
www.journalslibrary.nihr.ac.uk

*This report presents independent research funded by the National Institute for Health Research (NIHR).
The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the
Department of Health and Social Care*

Published by the NIHR Journals Library